

**IMPLEMENTASI ALGORITMA *RANDOM FOREST* UNTUK
MENGKLASIFIKASI HASIL PENGOBATAN KESEHATAN MENTAL**

SKRIPSI



Oleh:

SITI ROMLAH

2021502061

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS IBRAHIMY**

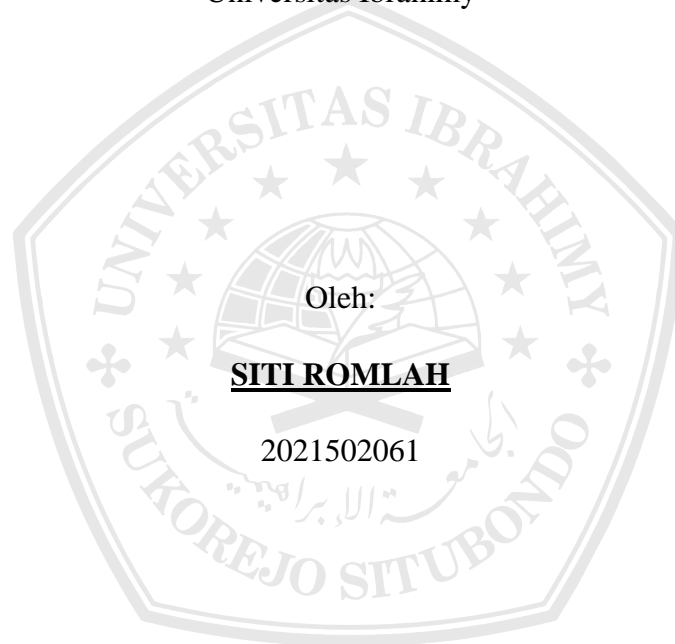
SITUBONDO

2025

**IMPLEMENTASI ALGORITMA *RANDOM FOREST* UNTUK
MENGKLASIFIKASI HASIL PENGOBATAN KESEHATAN MENTAL**

SKRIPSI

Diajukan untuk Memenuhi Salah Satu Persyaratan dalam Menyelesaikan Program Sarjana (S-1) pada Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Ibrahimy



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS IBRAHIMY
SITUBONDO**

2025

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : **Siti Romlah**

NPM/NRM : 2021502061

Program Studi : S-1 Sistem Informasi

Fakultas : Fakultas Sains dan Teknologi

Menyatakan dengan sebenarnya, bahwa skripsi ini secara keseluruhan adalah hasil penelitian atau karya saya sendiri, kecuali pada bagian-bagian yang dirujuk sebagai sumber referensi dan disebutkan dalam daftar pustaka. Apabila dikemudian hari terbukti atau dapat dibuktikan bahwa skripsi ini hasil plagiasi, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Situbondo, 19 Agustus 2025

Saya yang menyatakan,



Siti Romlah

PERSETUJUAN PEMBIMBING

Skripsi ditulis oleh:

Nama : **Siti Romlah**

NPM/NRM : 2021502061

Judul : **IMPLEMENTASI ALGORITMA *RANDOM FOREST*
UNTUK MENGLASIFIKASI HASIL PENGOBATAN
KESEHATAN MENTAL**

Telah ditelaah dan disetujui oleh pembimbing untuk diuji pada sidang/munaqosah.

Telah disetujui oleh:

Pembimbing I



Firman Santoso, M.Kom
NIDN: 0722129201

Pembimbing II



Nur Azise, M.Kom
NIDN : 0730108802

iii

iii

PENGESAHAN

SKRIPSI

IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK
MENGKLASIFIKASI HASIL PENGOBATAN KESEHATAN MENTAL

SITI ROMLAH
2021502061

Telah dipertanggung jawabkan di depan dewan penguji siding/munaqosah Skripsi pada hari Rabu, Tanggal 20 Agustus 2025 Sebagai salah satu syarat memperoleh gelar Sarjana (S.Kom) pada Fakultas Sains dan Teknologi Universitas Ibrahimi

Tim Penguji,

Ketua Sidang,

Abd. Ghofur, M.Kom
NIDN. 0711088303

Sekretaris Sidang,

Usman Hidayat, S.Kom
NIDN.

Penguji I,

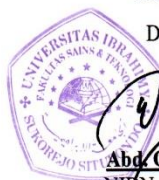
Adi Susanto, M.Kom
NIDN. 0708079104

Penguji II,

Lukman Fakhil Lidimillah, M.Kom
NIDN. 0715099001

Mengetahui

Dekan,



Abd. Ghofur, M.Kom
NIDN. 0711088303

MOTTO

“ BEKERJA DENGAN ILMU, BERJUANG DENGAN SABAR, DAN
BERTAWAKKAL KEPADA ALLAH ”



PERSEMBAHAN

Skripsi ini dipersembahkan teruntuk orang-orang yang telah berkontribusi besar dalam kehidupan peneliti serta memberi kontribusi dalam proses pembuatan Skripsi yang berjudul Implementasi Algoritma *Random Forest* Untuk Mengklasifikasi Hasil Pengobatan Kesehatan Mental ini:

1. Kedua orang tua yang saya cintai, dengan kasih sayang dan keikhlasan do'anya untuk sang buah hatinya di pondok, berjuang tanpa kenal lelah dan putus asa, karena beliaulah saya masih bisa bertahan demi menggapai impianku
2. Seluruh teman seperjuangan dan teman UKM saya khususnya prodi Sistem Informasi Angkatan 2021 yang selalu membantu dalam setiap kesulitan yang saya hadapi dalam menyelesaikan tugas selama ini
3. Teman-teman saya dikamar MQ.02 dan MQ.05 yang selalu sedia mendengarkan setiap cerita suka maupun duka
4. Semua orang yang tak bisa saya sebut satu persatu Namanya yang telah mendukung saya sepenuhnya sehingga saya dapat menyelesaikan laporan ini dengan sebaik baiknya.

KATA PENGANTAR

Segala puji syukur peneliti sampaikan kepada Allah SWT, karena atas Rahmat dan Hidayah-Nya, perencanaan, pelaksanaan dan penyelesaian skripsi dengan judul **“IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK MENGLASIFIKASI HASIL PENGobatan KESEHATAN MENTAL”** sebagai salah satu syarat penyelesaian program sarjana dapat terselesaikan dengan baik dan lancar.

Keberhasilan ini tentu tidak terlepas dari dukungan dan bimbingan berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, peneliti menyampaikan terima kasih yang sebesar-besarnya kepada:

1. KHR. Ach. Azaim Ibrahimi selaku pengasuh pondok pesantren Salafiyah Syafi'iyah
2. KH. Ach. Fadlail, SH.,M.H selaku rektor Universitas Ibrahimi
3. Bapak Abd. Ghofur, M.Kom selaku dekan Fakultas Sains dan Teknologi
4. Bapak Achmad Bajuri, M.Kom selaku Ketua Program Studi Sistem Informasi
5. Bapak Firman Santoso, M.Kom dan Ibu Nur Azise, M.Kom selaku pembimbing I dan II

Semoga semua amal baik yang telah diberikan oleh Bapak/Ibu kepada peneliti mendapat balasan yang sebaik mungkin dari Allah SWT, Amin.

Situbondo, 19 Agustus 2025

Peneliti

ABSTRAK

Siti Romlah. 2025. **Implementasi Algoritma Random Forest Untuk Mengklasifikasi Hasil Pengobatan Kesehatan Mental**. Skripsi, Program Studi Sistem Informasi, Universitas Ibrahimi. Pembimbing: (I) Firman Santoso, M.Kom, (II) Nur Azise, M.Kom.

Kesehatan mental menjadi isu penting akibat meningkatnya tekanan hidup dan pengaruh media sosial. Gangguan seperti depresi dan stress dapat menurunkan produktivitas serta kualitas hidup individu, sekaligus meningkatkan beban biaya layanan kesehatan mental hingga mencapai Rp87,5 triliun. Hal ini menunjukkan perlunya solusi inovatif dalam penanganannya. Penelitian ini bertujuan untuk mengimplementasikan algoritma *Random Forest* dalam mengklasifikasi hasil pengobatan kesehatan mental. Metode yang digunakan adalah eksperimen dengan pendekatan kuantitatif. Tahapan meliputi pengumpulan data, pre-processing data, pembagian data, penerapan algoritma, dan evaluasi menggunakan *confusion matrix*. Hasil penelitian menunjukkan bahwa metode *Random Forest* memiliki performa yang baik dengan akurasi sebesar 99,01%, sehingga dapat mendukung pengambilan keputusan dalam pengobatan kesehatan mental.

Kata Kunci: Kesehatan Mental, Klasifikasi, Random Forest, Pengobatan, Confusion Matrix.

ABSTRACT

Siti Romlah. 2025. **Implementation of Random Forest Algorithm to Classify Mental Health Treatment Outcomes**. Thesis, Information Systems Study Program, Ibrahimi University. Supervisors: (I) Firman Santoso, M.Kom, (II) Nur Azise, M.Kom.

Mental health has become a critical issue due to increasing life pressures and the influence of social media. Disorders such as depression and stress can reduce an individual's productivity and quality of life, while increasing the cost of mental health services to as much as IDR 87.5 trillion. This demonstrates the need for innovative solutions. This study aims to implement the Random Forest algorithm to classify mental health treatment outcomes. The method used is an experiment with a quantitative approach. The stages include data collection, data preprocessing, data sharing, algorithm implementation, and evaluation using a confusion matrix. The results show that the Random Forest method has good performance with an accuracy of 99.01%, thus supporting decision-making in mental health treatment.

Keywords: Mental Health, Classification, Random Forest, Treatment, Confusion Matrix.

DAFTAR ISI

| | Halaman |
|--|------------------------------|
| PERNYATAAN KEASLIAN TULISAN | Error! Bookmark not defined. |
| PERSETUJUAN PEMBIMBING | ii |
| PENGESAHAN SKRIPSI | iii |
| MOTTO | v |
| PERSEMBAHAN | vi |
| KATA PENGANTAR | vii |
| ABSTRAK | viii |
| ABSTRACT | ix |
| DAFTAR ISI | x |
| DAFTAR TABLE | xiii |
| DAFTAR GAMBAR | xiv |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang..... | 1 |
| 1.2 Identifikasi Masalah | 4 |
| 1.3 Rumusan Masalah | 5 |
| 1.4 Batasan Masalah..... | 5 |
| 1.5 Tujuan Penelitian..... | 5 |
| 1.6 Manfaat Penelitian..... | 5 |
| 1.7 Metode Penelitian..... | 6 |
| 1.7.1 Jenis Penelitian | 6 |
| 1.7.2 Teknik Pengumpulan Data..... | 6 |
| 1.8 Sistematika Pembahasan | 7 |
| BAB II TINJAUAN PUSTAKA | 9 |
| 2.1 Penelitain Terdahulu..... | 9 |
| 2.1.1 Klasifikasi Data Kesehatan Mental di Industri Teknologi Menggunakan Algoritma <i>Random Forest</i> | 9 |
| 2.1.2 Deteksi Tingkat Gangguan Kecemasan Menggunakan Metode <i>Random Forest</i> | 10 |
| 2.1.3 Diagnosa Tingkat Depresi Mahasiswa Selama Masa Pandemi <i>Covid-19</i> Menggunakan Algoritma <i>Random Forest</i> | 11 |

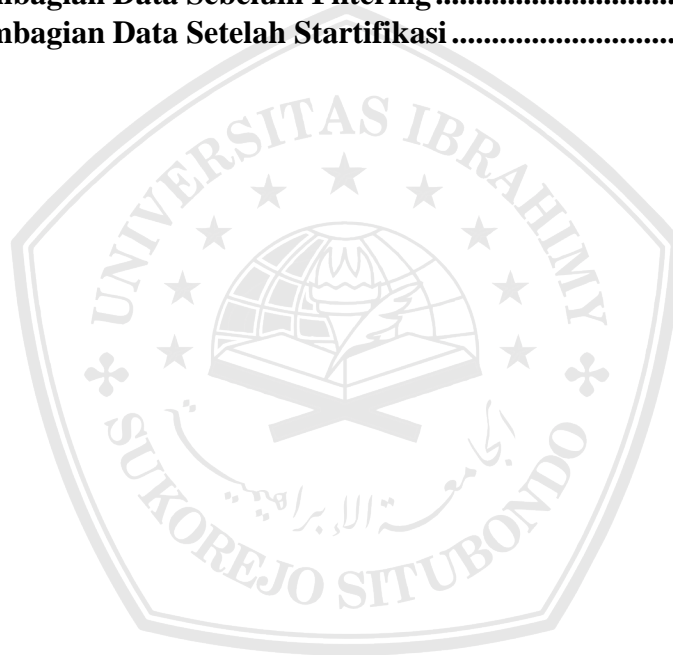
| | |
|---|-----------|
| 2.2 Landasan Teori | 12 |
| 2.2.1 Implementasi..... | 12 |
| 2.2.2 Klasifikasi | 12 |
| 2.2.3 Kesehatan Mental | 13 |
| 2.2.4 Data Mining | 13 |
| 2.2.5 Machine Learning..... | 14 |
| 2.2.6 Algoritma Random Forest | 15 |
| 2.3 Pemodelan | 19 |
| 2.3.1 Flowchart | 19 |
| 2.3.2 Data Flow Diagram (DFD) | 20 |
| 2.4 Perangkat Lunak Yang Digunakan..... | 22 |
| 2.4.1 Goggle Colab | 22 |
| 2.4.2 Microsoft Office Excel | 22 |
| 2.5 Bahasa Pemrograman | 22 |
| 2.5.1 Bahasa Pemrograman <i>Python</i> | 22 |
| BAB III ANALISIS DAN PERANCANGAN SISTEM..... | 24 |
| 3.1 Gambaran Umum Objek Penelitian..... | 24 |
| 3.2 Alur proses..... | 26 |
| 3.2.1 Identifikasi Dan Analisis Proses Bisnis | 26 |
| 3.2.2 Analisis Proses Bisnis..... | 27 |
| 3.2.3 Identifikasi Dan Analisis Kebutuhan..... | 32 |
| 3.2.4 Identifikasi Dan Analisis Alternatif Solusi | 34 |
| 3.3 Desain Proses..... | 35 |
| BAB IV IMPLEMENTASI SISTEM | 39 |
| 4.1 Konstruksi Sistem..... | 39 |
| 4.1.1 Kebutuhan Sistem..... | 39 |
| 4.2 Pengumpulan Data..... | 39 |
| 4.3 Pre-processing Data..... | 42 |
| 4.4 Pembagian Data (<i>Train-Test Split</i>)..... | 46 |
| 4.5 Metode Random Forest | 50 |
| BAB V PENUTUP..... | 61 |
| 5.1 Kesimpulan..... | 61 |

| | |
|------------------------------|-----------|
| 5.2 Saran..... | 62 |
| DAFTAR PUSTAKA..... | 63 |
| CURRICULUM VITAE..... | 66 |
| LAMPIRAN..... | 67 |



DAFTAR TABLE

| | |
|---|-----------|
| Tabel 2. 1 Confusion Matrix | 18 |
| Tabel 2. 2 Simbol Flowchart | 20 |
| Tabel 2. 3 Data Flow Diagram (DFD) | 21 |
| | |
| Tabel 3. 1 Identifikasi Alternatif Solusi | 34 |
| Tabel 3. 2 Analisis Kelayakan | 35 |
| | |
| Tabel 4. 1 Pembagian Data Sebelum Filtering | 48 |
| Tabel 4. 2 Pembagian Data Setelah Startifikasi | 49 |



DAFTAR GAMBAR

| | |
|---|-----------|
| Gambar 3. 1 Data file CSV..... | 25 |
| Gambar 3. 2 Tahap Pre-Processing..... | 28 |
| Gambar 3. 3 Flowchart Sistem | 31 |
| Gambar 3. 4 Pre-Processing Data..... | 36 |
| Gambar 3. 5 Metode Random Forest..... | 37 |
| Gambar 3. 6 Arsitektur Aplikasi | 38 |
| | |
| Gambar 4. 1 Dataset Mental Health Diagnosis And Treatment Monitoring | 41 |
| Gambar 4. 2 Distribusi Kelas Sebelum Filtering | 49 |
| Gambar 4. 3 Split Data | 50 |
| Gambar 4. 4 Confusion Matrix..... | 55 |



BAB I PENDAHULUAN

1.1 Latar Belakang

Pada era modern, kesadaran masyarakat terhadap pentingnya menjaga kesehatan mental mengalami peningkatan yang signifikan. Peningkatan ini dipengaruhi oleh berbagai faktor, seperti tekanan yang timbul dari lingkungan kerja, relasi sosial, tuntutan hidup yang kompleks, serta dampak negatif perkembangan teknologi dan media sosial, yang mendorong individu untuk lebih memperhatikan kondisi kesehatan mental mereka [1]. Menurut *World Health Organization* (WHO) kesehatan mental adalah kondisi kesejahteraan (*well-being*) dimana individu mampu menyadari potensi dirinya, mengatasi tekanan hidup yang normal, bekerja secara produktif, serta berkontribusi bagi komunitasnya [2]. Gangguan suasana hati, khususnya depresi, sering kali ditandai dengan gejala seperti kelelahan, hilangnya ketertarikan terhadap aktivitas, rasa bersalah yang mendalam, kesulitan dalam fokus, penurunan nafsu makan, serta munculnya ide atau keinginan untuk mengakhiri hidup, yang merupakan bagian dari manifestasi klinis gangguan kesehatan mental [3].

Sebagai negara dengan jumlah penduduk yang tinggi dan latar belakang budaya yang beragam, Indonesia mulai memberikan perhatian lebih terhadap permasalahan kesehatan mental. Isu ini kini telah dipandang sebagai elemen krusial dalam kebijakan dan program kesehatan masyarakat [4]. Masyarakat perlu membangun ketahanan psikologis serta mengembangkan strategi untuk menghadapi berbagai tantangan kesehatan mental secara menyeluruh.

Tingginya tingkat stress, gangguan suasana hati, dan berbagai gangguan mental lainnya berdampak luas pada individu maupun masyarakat. Dampak tersebut mencakup penurunan produktivitas, peningkatan beban biaya layanan kesehatan, hingga menurunnya kualitas hidup. Berdasarkan data tahun 2023 dan 2024, tercatat bahwa biaya langsung penanganan gangguan jiwa di Indonesia mencapai Rp87,5 triliun. Fakta ini menegaskan bahwa permasalahan kesehatan mental tidak hanya berdampak pada aspek medis, tetapi juga ekonomi. Oleh karena itu, dibutuhkan kebijakan yang menyeluruh serta kolaborasi lintas sektor untuk meningkatkan kesadaran, pencegahan, dan akses terhadap layanan kesehatan jiwa.

Kesehatan mental menggambarkan keadaan di mana aspek kepribadian, emosi, intelektual, dan fisik seseorang dapat berfungsi secara maksimal. Individu dengan kondisi mental yang sehat mampu menyesuaikan diri terhadap lingkungan dan tekanan hidup, menjalankan perannya secara seimbang, memiliki penerimaan diri yang baik, serta mampu menghadapi tantangan hidup dengan rasa puas dan kebahagiaan [5].

Perkembangan teknologi, khususnya di bidang kecerdasan buatan (*Artificial Intelligence*), membuka peluang baru dalam mendukung proses diagnosis dan pengobatan gangguan kesehatan mental melalui pendekatan prediktif. Salah satu algoritma yang dikenal memiliki performa baik dalam klasifikasi adalah *Random Forest*. Algoritma ini mampu mengelola data dalam jumlah besar dan kompleks secara efektif, meskipun masih menghadapi kendala seperti risiko *overfitting*, terutama ketika menangani data yang tidak seimbang [6].

Penelitian yang dilakukan oleh Agus Priyono, Muhammad Shodiq, Dwi Putra Alvinsyah, Septina Alfiani Hidayah (2024). Mengenai “Metode *Random Forest* Untuk Memudahkan Klasifikasi Diagnosis Penyakit Mental”. menunjukkan bahwa *Random Forest* efektif untuk mendiagnosis gangguan mental dengan akurasi tinggi. Penelitian ini menggunakan dataset dari Harvard University yang berisi 120 data pasien dengan 17 atribut, termasuk 14 data kategori yang sudah melalui *encoding*. Data dilabel dalam empat kategori: depresi, bipolar tipe 1, bipolar tipe 2, dan normal. Setelah dinormalisasi dengan *Min-Max Normalization*, hasilnya menunjukkan akurasi 90,83%, presisi 93,25%, dan recall 90,83%, yang mengindikasikan potensi *Random Forest* dalam mendukung diagnosis gangguan mental dan pengambilan keputusan medis yang lebih efisien [7].

Penelitian yang dilakukan oleh Sza Sza Amulya Larasati, Elok Nuraida Kusuma Dewi, Brahma Hanif Farhansyah, Fitra Abdurrachman Bachtiar, Fajar Pradana (2024). Mengenai “Penerapan *Decision Tree* Dan *Random Forest* Dalam Deteksi Tingkat Stress Manusia Berdasarkan Kondisi Tidur”. menunjukkan bahwa *Random Forest* dinilai lebih unggul dalam mendeteksi tingkat stres manusia berdasarkan kondisi tidur karena lebih akurat dan mampu mengurangi misklasifikasi yang terjadi pada model *Decision Tree*. Dua algoritma klasifikasi yaitu *Decision Tree* dan *Random Forest* digunakan dalam pengembangan model, dengan proses normalisasi data, *cross-validation*, serta *hyperparameter tuning*. Data yang digunakan memiliki 630 baris dengan 8 fitur fisiologis. Hasilnya, model *Decision Tree* mencapai akurasi 0,99, sementara *Random Forest* mencapai akurasi

sempurna 1,00, menunjukkan bahwa *Random Forest* lebih efektif dalam klasifikasi tingkat stres berdasarkan kondisi tidur [8].

Berdasarkan uraian diatas, penelitian ini bertujuan untuk menerapkan algoritma *Random Forest* dalam mengklasifikasi hasil pengobatan kesehatan mental. Diharapkan, metode ini dapat mendukung tenaga kesehatan dan pasien dalam pengambilan keputusan medis yang lebih akurat dan cepat. Selain itu, penelitian ini juga diharapkan dapat memberikan kontribusi dalam pengembangan sistem prediksi berbasis kecerdasan buatan untuk pengelolaan kesehatan mental yang lebih efektif dan efisien.

1.2 Identifikasi Masalah

Berdasarkan latar belakang penelitian ini terdapat beberapa identifikasi permasalahan sebagai berikut :

- a. Banyaknya orang yang tidak menyadari bahwa mereka memiliki gangguan kesehatan mental seperti tingginya *stress*, kecemasan, depresi dan gangguan kesehatan mental lainnya.
- b. Banyaknya biaya kesehatan mental yang terus meningkat setiap tahunnya.
- c. Dibutuhkan metode yang tepat untuk memprediksi hasil pengobatan kesehatan mental.
- d. Dibutuhkan optimasi pada algoritma *Random Forest* dalam memprediksi hasil pengobatan kesehatan mental.

1.3 Rumusan Masalah

Berdasarkan latar belakang dan identifikasi masalah di atas dapat ditentukan rumusan masalah dalam penelitian ini adalah bagaimana implementasi algoritma *Random Forest* digunakan untuk mengklasifikasi hasil pengobatan kesehatan mental?

1.4 Batasan Masalah

Dari penjelasan diatas dapat dijadikan landasan untuk menentukan batasan masalah dalam penelitian yang dilakukan. Adapun batasan masalah yang ada pada penelitian ini adalah :

- a. Data yang digunakan berasal dari dataset terpublikasi dari platform *Kaggle* tentang diagnosis dan pemantauan perawatan kesehatan mental.
- b. Algoritma yang digunakan adalah algoritma *Random Forest* dalam mengklasifikasi hasil pengobatan kesehatan mental

1.5 Tujuan Penelitian

Tujuan dari penelitian ini adalah :

- a. Mengembangkan model klasifikasi berbasis *Random Forest* untuk diagnosis dini gangguan kesehatan mental.
- b. Memberikan kontribusi pada pengembangan teknologi berbasis bukti untuk mendukung kebijakan dan pengelolaan kesehatan mental secara lebih efektif.

1.6 Manfaat Penelitian

Manfaat penelitian berdasarkan latar belakang diatas adalah sebagai berikut :

- a. Memberikan pemahaman lebih mendalam tentang penerapan algoritma *Random Forest* dalam diagnosis dan klasifikasi hasil pengobatan kesehatan mental.
- b. Mendukung deteksi dini gangguan kesehatan mental untuk mengurangi dampak negatif pada produktivitas dan kualitas hidup.
- c. Memberikan data berbasis bukti untuk mendukung kebijakan kesehatan mental yang lebih terarah dan efektif.

1.7 Metode Penelitian

1.7.1 Jenis Penelitian

Penelitian ini menggunakan pendekatan eksperimen, yakni suatu metode yang bertujuan untuk mengkaji pengaruh suatu perlakuan terhadap variabel lain dalam kondisi yang dikendalikan. Metode ini juga dikenal sebagai penelitian eksperimental, yang bersifat prediktif karena bertujuan memperkirakan dampak dari manipulasi variabel bebas terhadap variabel terikat [9].

Pada penelitian ini dilakukan dengan cara menerapkan *algoritma Random Forest* pada dataset pasien yang menjalani pengobatan kesehatan mental, kemudian membandingkan hasil prediksi model dengan data aktual. Evaluasi dilakukan menggunakan metrik performa seperti *akurasi*, *precision*, *recall*, dan *F1-score* untuk menilai keakuratan prediksi guna menentukan algoritma yang paling optimal dalam memprediksi hasil pengobatan pasien.

1.7.2 Teknik Pengumpulan Data

Data yang digunakan dalam penelitian ini berasal dari forum *Kaggle*, forum *Kaggle* ini berisi semua dataset public. *Kaggle* adalah platform yang

menyediakan berbagai dataset untuk kompetisi pembelajaran mesin dan penelitian [10].

1.8 Sistematika Pembahasan

Sistematika pembahasan yang digunakan dalam penelitian ini yaitu adalah sebagai berikut :

Bab I : Pendahuluan

Pada bab ini menjelaskan tentang latar belakang, identifikasi masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metode penelitian, sistematika pembahasan mengenai objek yang sedang diteliti.

Bab II : Tinjauan Pustaka

Pada bab ini menjelaskan tentang tinjauan pustaka dari adanya penelitian ini. Tinjauan pustaka berisi referensi seperti jurnal, skripsi, atau sebagainya yang berfungsi sebagai penunjang dari adanya penelitian terhadap objek yang diteliti. Terdapat beberapa komponen pada bab ini yang terdiri dari penelitian terdahulu, landasan teori, pemodelan, dan perangkat lunak yang digunakan.

Bab III : Analisis Perancangan Sistem

Pada bab ini menjelaskan tentang gambaran alur proses dan desain dari objek penelitian yang diteliti, terdapat beberapa sub bab pada bagian ini yakni identifikasi dan analisis proses bisnis, identifikasi dan analisis kebutuhan, identifikasi dan alternatif solusi, sedangkan desain sistem membahas tentang perancangan alur sistem.

Bab IV : Implementasi Sistem

Pada bagian ini menjelaskan tentang alur dari perhitungan algoritma, pengujian algoritma.

Bab V : Penutup

Pada bab ini menjelaskan tentang kesimpulan dan hasil dari penelitian yang dilakukan.



BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

2.1.1 Klasifikasi Data Kesehatan Mental di Industri Teknologi Menggunakan Algoritma *Random Forest*.

Dalam penelitian yang dilakukan oleh Emia Rosta Br. Sebayang, Yulison Herry Chrisnanto, Melina (2023). Penelitian ini bertujuan untuk mengklasifikasikan data kesehatan mental karyawan di industri teknologi menggunakan algoritma *Random Forest*. Metode yang digunakan mencakup pemanfaatan dataset survei kesehatan mental dari *Open Source Mental Illness* (OSMI), penerapan algoritma *Random Forest* dengan metode *ensemble* dan *bootstrap tree*, serta pembagian data menjadi training data (80%) dan test data (20%). Evaluasi model dilakukan menggunakan *Confusion Matrix* untuk mengukur akurasi, presisi, recall, dan *F1-score*. Dalam penelitian ini, terdapat 15 atribut yang digunakan untuk klasifikasi kesehatan mental, yaitu usia, jenis kelamin, riwayat keluarga, pengobatan, gangguan kerja, kerja jarak jauh, perusahaan teknologi, pilihan perawatan, program kesehatan, mencari bantuan, anonimitas, konsekuensi kesehatan mental, *supervisor*, wawancara kesehatan mental, dan wawancara kesehatan fisik.

Hasil penelitian menunjukkan bahwa algoritma *Random Forest* memiliki tingkat akurasi yang cukup baik dalam mengklasifikasikan data kesehatan mental, dengan nilai tertinggi mencapai 84,12%. Model dengan proporsi data latih 80% dan

data uji 20% menghasilkan performa terbaik dibandingkan skenario lainnya. Temuan ini menunjukkan bahwa semakin banyak data yang digunakan dalam pelatihan, semakin baik model dalam mengklasifikasikan kondisi kesehatan mental karyawan. Dengan hasil yang diperoleh, penelitian ini diharapkan dapat berkontribusi dalam meningkatkan kesadaran dan perhatian perusahaan terhadap kesehatan mental karyawan, sehingga berdampak positif pada kualitas hidup serta produktivitas di lingkungan kerja industri teknologi [6].

2.1.2 Deteksi Tingkat Gangguan Kecemasan Menggunakan Metode *Random Forest*.

Dalam penelitian yang dilakukan oleh Ghelifira Nur Fadhilah, Rizal Adi Saputra, Asa Hari Wibowo (2024). Penelitian ini bertujuan untuk mengidentifikasi tahap awal gangguan kecemasan menggunakan Data Mining dengan algoritma *Random Forest*. Data dikumpulkan melalui kuesioner GAD-7 (*General Anxiety Disorder-7*) yang diisi oleh 670 responden. Proses pengolahan data mengikuti metode *Knowledge Discovery in Database* (KDD), yang mencakup lima tahapan utama: data *selecting*, data *preprocessing*, data *transformation*, data mining, dan *evaluation*. Pada tahap data *selecting*, dipilih data yang relevan, sementara tahap data *preprocessing* melibatkan penghapusan atribut yang tidak diperlukan untuk meningkatkan kualitas data. Selanjutnya, data *transformation* mengubah data nominal menjadi numerik berdasarkan skor GAD-7. Proses data mining dilakukan dengan menerapkan algoritma *Random Forest* untuk mengklasifikasikan tingkat kecemasan, dan model yang dihasilkan dievaluasi menggunakan metode *10-fold cross-validation* dengan pengukuran *specificity*, *sensitivity*, dan *accuracy*.

Penelitian ini menggunakan tujuh atribut utama dari *kuesioner GAD-7*, yaitu perasaan gugup atau cemas, ketidakmampuan mengontrol kekhawatiran, kekhawatiran berlebihan, kesulitan merasa santai, perasaan gelisah, mudah tersinggung, dan ketakutan berlebihan terhadap kemungkinan hal buruk terjadi. Hasil analisis menunjukkan bahwa model *Random Forest* memiliki akurasi 89,55% dalam mendeteksi tingkat kecemasan. Dari 670 responden, 12,99% tergolong normal, 28,06% mengalami kecemasan ringan, 32,84% mengalami kecemasan sedang, dan 26,12% mengalami kecemasan berat. Model ini menunjukkan performa yang baik dengan sensitivity di atas 87% untuk setiap tingkat kecemasan, sehingga berpotensi digunakan sebagai alat bantu dalam deteksi dini gangguan kecemasan [11]

2.1.3 Diagnosa Tingkat Depresi Mahasiswa Selama Masa Pandemi *Covid-19* Menggunakan Algoritma *Random Forest*.

Penelitian yang dilakukan oleh Dewi Septiani, Ultach Enri, Nina Sulistiyowati (2021). Penelitian ini bertujuan untuk mendiagnosis tingkat depresi mahasiswa Universitas Singaperbangsa Karawang selama pandemi *COVID-19* menggunakan teknik data mining dengan algoritma *Random Forest*. Metode yang diterapkan adalah *Knowledge Discovery in Database (KDD)*, yang mencakup tahapan data *selection*, data *preprocessing*, data *transformation*, data mining, dan *evaluation*. Data dikumpulkan melalui *kuesioner PHQ-9 (Patient Health Questionnaire-9)* yang diisi oleh 392 responden. Kuesioner ini terdiri dari sembilan atribut utama yang mencerminkan gejala depresi, seperti kurangnya minat, perasaan murung, gangguan tidur, kelelahan, perubahan nafsu makan, serta pemikiran untuk

melukai diri sendiri. Model divalidasi menggunakan *10-fold cross-validation* dengan mengukur *accuracy*, *sensitivity*, dan *specificity*.

Hasil penelitian menunjukkan bahwa model *Random Forest* memiliki akurasi sebesar 85,94% dalam mendeteksi tingkat depresi. Dari total responden, 1,02% tergolong normal, 47,96% mengalami gejala depresi ringan, 36,73% mengalami depresi ringan, 8,16% mengalami depresi sedang, dan 6,12% mengalami depresi berat. Berdasarkan hasil ini, 51,79% mahasiswa disarankan untuk menjalani pemeriksaan lebih lanjut ke psikolog. Penelitian ini membuktikan bahwa algoritma *Random Forest* dapat digunakan sebagai alat skrining awal yang efektif dalam mendeteksi tingkat depresi mahasiswa, sehingga dapat membantu pengambilan keputusan dalam intervensi kesehatan mental [12].

2.2 Landasan Teori

Berikut adalah landasan yang mendukung adanya penelitian ini antara lain :

2.2.1 Implementasi

Implementasi dapat dipahami sebagai tahapan pelaksanaan di mana suatu rencana, metode, konsep, atau kebijakan dijalankan dalam bentuk tindakan konkret untuk mewujudkan tujuan yang telah dirumuskan sebelumnya [13].

2.2.2 Klasifikasi

Klasifikasi merupakan salah satu metode analisis data yang bertujuan membangun model guna menggambarkan peranan suatu kelas data. Dalam implementasinya, metode ini digunakan untuk memperkirakan label kategori yang bersifat diskrit serta tidak memiliki urutan. Melalui pendekatan tersebut, peneliti

dapat memperoleh pemahaman yang lebih menyeluruh terhadap data yang dianalisis [14].

2.2.3 Kesehatan Mental

Kesehatan mental dapat diartikan sebagai kondisi di mana individu terbebas dari gejala gangguan maupun penyakit kejiwaan, mampu beradaptasi dengan baik, serta dapat mengembangkan potensi dan bakat secara optimal. Kondisi ini turut mendukung terciptanya kebahagiaan bersama dan keharmonisan batin dalam menjalani kehidupan [15].

Faktor penting dalam upaya peningkatan kesehatan mental terletak pada sikap-sikap yang dimiliki oleh individu maupun kelompok sosial tempat individu tersebut bernaung. Secara umum, sikap yang mendukung kesehatan mental mencakup penghargaan terhadap diri sendiri, penerimaan terhadap keterbatasan pribadi maupun orang lain, pemahaman bahwa setiap perilaku memiliki latar belakang tertentu, serta kesadaran akan dorongan untuk mengaktualisasikan diri [16].

2.2.4 Data Mining

Data mining merupakan proses untuk mengeksplorasi dan menganalisis pola-pola tersembunyi dalam data dari berbagai sudut pandang, dengan tujuan mengklasifikasikan data menjadi informasi yang bermanfaat. Informasi ini biasanya dikumpulkan dalam area penyimpanan terpusat seperti data *warehouse* guna mendukung analisis yang lebih efisien. Melalui penerapan algoritma *data mining*, proses ini dapat membantu dalam pengambilan keputusan bisnis serta penyediaan informasi penting lainnya [17].

Data mining, yang juga dikenal sebagai penemuan pengetahuan atau *knowledge discovery*, merupakan proses untuk mengidentifikasi pola-pola tersembunyi dalam himpunan data. Dalam praktiknya, data mining mengekstraksi informasi yang bernilai melalui analisis terhadap pola dan hubungan tertentu yang terdapat dalam kumpulan data berukuran besar [18].

Dalam konteks penelitian ini, *data mining* digunakan untuk membantu penelitian prediksi hasil pengobatan kesehatan mental dengan menganalisis data pasien untuk menemukan pola, tren, dan faktor yang memengaruhi keberhasilan terapi dengan menggunakan algoritma *Random Forest*. *Supervised learning* adalah salah satu metode *data mining* khususnya dalam kategori *predictive analytics* yang bertujuan untuk mempelajari hubungan antara data masukan (fitur pasien) dan hasil pengobatan (label), sehingga model dapat membuat prediksi untuk pasien baru. Dengan menerapkan algoritma *Random Forest* pada data diagnosis dan pemantauan perawatan kesehatan mental yang diambil dari *Kaggle*. Penelitian ini bertujuan untuk mengklasifikasi hasil pengobatan kesehatan mental sehingga dapat mendukung pengambilan keputusan yang lebih tepat dan meningkatkan hasil terapi untuk pasien.

2.2.5 Machine Learning

Machine Learning atau pembelajaran mesin merupakan cabang dari kecerdasan buatan (*Artificial Intelligence*) yang berfokus pada pengembangan algoritma dan model statistik, yang memungkinkan sistem komputer untuk meningkatkan performa dalam menyelesaikan tugas tertentu seiring waktu tanpa perlu instruksi pemrograman secara langsung [19].

Dengan memanfaatkan *Machine Learning*, khususnya algoritma *Random Forest*, dapat diterapkan untuk mengklasifikasi hasil pengobatan kesehatan mental. Dengan memanfaatkan data historis pasien, model prediktif dapat memberikan wawasan yang lebih baik tentang efektivitas pengobatan, membantu tenaga medis untuk membuat keputusan yang lebih berbasis data.

2.2.6 Algoritma Random Forest

Secara prinsip, *Random Forest* merupakan pengembangan lanjutan dari algoritma *Decision Tree*. Oleh karena itu, jika seseorang telah memahami cara kerja dan dasar pembentukan *Decision Tree*, maka proses memahami *Random Forest* akan menjadi relatif lebih sederhana [20].

Random Forest merupakan metode klasifikasi berbasis *ensemble learning* yang menggabungkan sejumlah pohon keputusan. Setiap pohon dibentuk menggunakan sampel data dan fitur yang dipilih secara acak, dengan tujuan menghasilkan prediksi kelas yang lebih akurat dan stabil [21].

Berikut adalah langkah-langkah perhitungan dalam algoritma *Random Forest* sebagai berikut :

1. *Bootstrap Sampling*

Dari dataset asli berisi N data, algoritma memilih secara acak beberapa subset data dengan penggantian dari data set asli.

Rumus :

$$D_i = x_j, y_j, j \in 1, 2, \dots, n \quad (2.1)$$

2. Pembentukan pohon keputusan

Untuk setiap subset data, sebuah pohon keputusan dibangun. Pada setiap node dalam pohon, algoritma memilih secara acak sejumlah fitur (biasanya \sqrt{M} dari total M fitur) dan menentukan pemisahan terbaik berdasarkan fitur-fitur tersebut. Rumus :

$$M = \sqrt{p} \text{ (untuk klasifikasi)} \quad (2.2)$$

$$M = \frac{p}{3} \text{ (untuk regresi, aturan empiris)}$$

3. Pembentukan hutan

Proses pada langkah kedua diulang untuk membentuk sejumlah pohon keputusan, sehingga membentuk suatu "hutan" dari pohon-pohon keputusan.

Dari subset fitur m yang dipilih, cari fitur terbaik untuk melakukan split menggunakan kriteria seperti :

a. *Gini Impurity*

Gini impurity mengukur probabilitas bahwa suatu data yang dipilih secara acak akan salah diklasifikasikan jika diberi label secara acak berdasarkan distribusi kelas yang ada dalam node tersebut. Rumusnya adalah :

$$Gini = 1 - \sum_{i=1}^n P_i^2 \quad (2.3)$$

Di mana P_i adalah proporsi sampel pada kelas i di node tersebut. Semakin kecil nilai *Gini impurity*, semakin baik atribut tersebut dalam membagi data.

b. *Entropy dan Information Gain.*

Entropy mengukur tingkat ketidakteraturan dalam data, dengan rumus :

$$Entropy = \sum_{i=1}^n P_i \log_2 P_i \quad (2.4)$$

Semakin tinggi *entropi*, semakin beragam kelas dalam node tersebut.

Information Gain digunakan untuk memilih atribut terbaik dalam pemisahan node, di mana semakin besar *Information Gain*, semakin baik atribut tersebut dalam mengurangi *impurity*.

c. Ulangi proses pembagian sampai salah satu dari kondisi berikut tercapai:

- a) Node hanya memiliki satu kelas.
- b) Node tidak memiliki fitur yang dapat digunakan untuk split.
- c) Pohon mencapai kedalaman maksimum yang ditentukan.

4. Prediksi

Untuk data baru, setiap pohon dalam hutan memberikan prediksi. Dalam kasus klasifikasi, hasil akhir ditentukan berdasarkan mayoritas suara (voting) dari semua pohon [22]. Dengan rumus :

Untuk rata-rata prediksi :

$$\hat{y} = \frac{1}{m} \sum_{i=1}^m \hat{y}_i \quad (2.5)$$

5. Evaluasi model

Proses evaluasi digunakan untuk mengetahui nilai kinerja algoritma *Random Forest* dalam memprediksi hasil pengobatan kesehatan mental seperti *accuracy*, *precision*, *recall* dan *F1-Score* yang dinyatakan pada table *confusion matrix* berikut ini :

Tabel 2. 1 Confusion Matrix

| | Prediksi Positive | Prediksi Negative |
|------------------------|----------------------------|----------------------------|
| Aktual Positive | <i>True Positive (TP)</i> | <i>False Negative (FN)</i> |
| Aktual Negative | <i>False Positive (FP)</i> | <i>True Negative (TN)</i> |

Analisis dari prediksi hasil pengobatan kesehatan mental dengan *confusion matrix* memberikan gambaran yang jelas tentang kinerja model dalam memprediksi hasil pengobatan yaitu :

- a. *True Positive (TP)* berarti pasien yang diprediksi akan berhasil dan memang berhasil dalam pengobatan.
- b. *False Positive (FP)* berarti pasien yang diprediksi akan berhasil, tetapi sebenarnya gagal.
- c. *True Negative (TN)* berarti pasien yang diprediksi akan gagal dan memang gagal.
- d. *False Negative (FN)* berarti pasien yang diprediksi akan gagal, tetapi sebenarnya berhasil.

Berdasarkan rumus perhitungan sebagai berikut :

- a. Akurasi (*Accuracy*)

Persentase prediksi yang benar dari total prediksi

$$Accuracy = \frac{\text{Jumlah prediksi benar (Positif dan negatif)}}{\text{Jumlah keseluruhan data}} \quad (2.6)$$

Akurasi tinggi menunjukkan model mampu memprediksi mayoritas hasil pengobatan dengan benar.

- b. Presisi (*Precision*)

Mengukur akurasi prediksi positif yang benar terhadap total prediksi positif.

$$Precision = \frac{TP(TruePositive)}{TP(TruePositif) + FP(FalsePositive)} \quad (2.7)$$

presisi penting dalam konteks di mana prediksi positif harus sangat andal, seperti memastikan pasien menerima pengobatan yang benar.

c. Recall (*Sensitivity*)

Mengukur kemampuan model untuk menangkap semua kasus positif.

$$Recall = \frac{TP(TruePositive)}{TP(TruePositif) + FN(FalseNegative)} \quad (2.8)$$

recall relevan jika prioritas adalah meminimalkan kesalahan dalam mendeteksi kasus positif (misalnya, pasien yang membutuhkan perhatian lebih).

d. F1-Score

Rata-rata harmonis dari *presisi* dan *recall*, memberikan keseimbangan antara keduanya :

$$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (2.9)$$

Berguna untuk menyeimbangkan *precision* dan *recall*, terutama saat kelas target tidak seimbang.





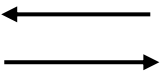

2.3 Pemodelan

2.3.1 Flowchart

Flowchart atau diagram alir merupakan representasi visual dari suatu algoritma, proses, atau alur kerja, yang menyajikan setiap langkah menggunakan simbol-simbol grafis yang dihubungkan oleh panah sebagai penanda urutan. *Flowchart* digunakan sebagai alat dokumentasi untuk menggambarkan proses dan

tahapan yang telah dirancang. Karena disusun secara visual, flowchart memiliki kelebihan dalam menjelaskan algoritma secara lebih jelas dan terstruktur. Oleh sebab itu, dalam standar internasional, algoritma lebih umum disajikan secara visual melalui *flowchart* menggunakan simbol-simbol standar, bukan dalam bentuk narasi teks [23]. Simbol-simbol *Flowchart* sebagaimana yang dijelaskan dalam table 2.2 berikut:

Tabel 2.2 Simbol Flowchart

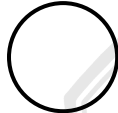
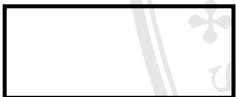

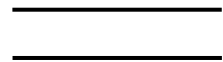
| NO | SIMBOL | NAMA | FUNGSI |
|----|---|----------------------------------|--|
| 1. |  | Terminator | Menunjukkan awal atau akhir proses. |
| 2. |  | Proses | Menunjukkan aktivitas atau langkah yang dilakukan dalam suatu proses. |
| 3. |  | Keputusan (<i>Decision</i>) | Menyatakan titik pengambilan keputusan dengan beberapa kemungkinan hasil (contoh: Ya/Tidak). |
| 4. |  | <i>Input/output</i> | Menunjukkan aktivitas yang menerima input atau menghasilkan output dari suatu sistem. |
| 5. |  | Aliran (<i>Flowline</i>) | Menunjukkan arah aliran data atau urutan langkah proses dari satu elemen ke elemen lainnya. |
| 6. |  | <i>Manual Operation</i> | Menunjukkan langkah yang dilakukan secara manual tanpa sistem otomatis. |

2.3.2 Data Flow Diagram (DFD)

Data Flow Diagram (DFD) adalah suatu bentuk visualisasi yang menggambarkan aliran data dalam sebuah sistem melalui penggunaan simbol-

simbol khusus. Diagram ini sangat bermanfaat dalam memahami sistem secara logis, terstruktur, dan jelas. DFD juga menunjukkan bagaimana data diolah oleh sistem, mulai dari tahap masukan (*input*) hingga menjadi keluaran (*output*), dengan penekanan pada aliran informasi hasil pengobatan kesehatan mental meliputi asal data, tujuan akhir, serta cara penyimpanannya [24]. Sebagaimana table *Data Flow Diagram* (DFD) yang dijelaskan dalam table 2.3 berikut :

Tabel 2. 3 Data Flow Diagram (DFD)

| NO | SIMBOL | NAMA | KETERANGAN |
|----|---|------------------------------------|--|
| 1. |  | Proses (<i>Order Processing</i>) | menunjukkan aktivitas atau fungsi utama dalam sistem yang mengolah data dari entitas eksternal atau penyimpanan data. |
| 2. |  | Entitas Eksternal | Mewakili aktor atau sistem di luar sistem yang sedang dimodelkan. Mereka adalah sumber atau penerima data dari sistem utama. |
| 3. |  | Aliran Data | Menunjukkan pergerakan data dari satu elemen ke elemen lain, baik dari entitas eksternal ke proses, antar proses, atau dari/ke penyimpanan data. |
| 4. |  | Penyimpanan Data | Mewakili lokasi penyimpanan data yang digunakan untuk menyimpan informasi sementara atau permanen selama sistem beroperasi. |

2.4 Perangkat Lunak Yang Digunakan

2.4.1 Goggle Colab

Google Colaboratory (Google Colab) atau dikenal juga sebagai *interactive notebook* milik *Google*, merupakan lingkungan pemrograman berbasis *Python* yang menggunakan format '*notebook*', serupa dengan *Jupyter Notebook*. Secara sederhana, *Google* menyediakan akses komputasi secara gratis untuk menjalankan program *Python*. Melalui *Google Colab*, pengguna dapat mengeksekusi kode *Python* langsung dari *browser*. Selain itu, *platform* ini juga memungkinkan pengguna untuk menulis, menyimpan, serta membagikan kode program yang telah dibuat melalui integrasi dengan *Google Drive* [25].

2.4.2 Microsoft Office Excel

Microsoft Office Excel merupakan program pengolah *spreadsheet* yang memungkinkan pengguna untuk memanipulasi data dalam bentuk teks, numerik, maupun alfanumerik, yang umumnya tersusun dalam tabel. Hasil pengolahan data ini dapat digunakan sebagai informasi pendukung dalam proses pengambilan keputusan, baik dalam konteks permasalahan sederhana maupun yang lebih kompleks [26].

2.5 Bahasa Pemrograman

2.5.1 Bahasa Pemrograman *Python*

Python merupakan bahasa pemrograman tingkat tinggi (*high-level programming language*) yang bersifat *interpreted* dan dapat digunakan untuk berbagai kebutuhan pemrograman (*general-purpose*). Istilah 'tingkat tinggi'

mengacu pada tingkat kemiripan sintaks bahasa pemrograman dengan bahasa manusia, yang membuatnya lebih mudah dipahami dibandingkan dengan bahasa mesin yang digunakan komputer. Sebagai bahasa *interpreted*, *Python* mengeksekusi kode secara langsung baris demi baris tanpa melalui proses kompilasi terlebih dahulu. Mekanisme ini serupa dengan cara kerja bahasa pemrograman *skrip* seperti PHP dan *JavaScript* [27]. beberapa fitur utama *python* adalah sebagai berikut :

1. Sintaksis yang sederhana
2. Dukungan untuk berbagai platform
3. Berbasis interpreter
4. Dukungan untuk berbagai paradigma pemrograman
5. Dinamis dan tipe data yang kuat
6. Memiliki berbagai *library* dan *framework*
7. Mudah diperluas
8. Komunitas yang besar

Python digunakan secara luas di berbagai sektor, termasuk pengembangan web, analisis data, kecerdasan buatan, dan pembuatan perangkat lunak. Kemampuannya yang tinggi dalam hal keterbacaan kode serta fleksibilitas menjadikan *Python* sebagai salah satu bahasa pemrograman yang paling populer di kalangan *developer* [28].

BAB III

ANALISIS DAN PERANCANGAN SISTEM

3.1 Gambaran Umum Objek Penelitian

Kesehatan mental merupakan kondisi kesejahteraan emosional, psikologis, dan sosial yang memengaruhi cara individu berpikir, merasa, dan bertindak. Gangguan mental seperti depresi dan kecemasan terus meningkat akibat tekanan hidup, konflik sosial, dan dampak negatif media sosial. Di Indonesia, angka kejadian gangguan mental terus mengalami peningkatan yang mengkhawatirkan dan memberikan dampak negatif terhadap tingkat produktivitas serta menambah beban ekonomi.. Dataset *Mental Health Diagnosis And Treatment Monitoring* merupakan data sintesis dari *Kaggle* yang berisi 500 data pasien, yang mencakup usia, jenis kelamin, tingkat keparahan gejala, diagnosis, jenis terapi, medikasi, serta hasil pengobatan. Dengan mengolah data ini menggunakan algoritma *Random Forest*, dapat dibangun model prediksi keberhasilan pengobatan yang bermanfaat bagi praktisi dan lembaga kesehatan. Berikut adalah tahapan proses pemanfaatan dataset:

1. Mengakses dataset *Mental Health Diagnosis And Treatment Monitoring*.
2. Mengunduh file dataset dalam bentuk CSV.

| Patient ID | Age | Gender | Diagnosis | Symptom Severity (1-10) | Mood Score (1-10) | Sleep Quality (1-10) | Physical Activity (hrs/week) | Medication | Therapy Type | Treatment Start Date | Treatment Duration (weeks) | Stress Level (1-10) | Treatment Progress (1-10) | AI-Detected Emotional State | Adherence to Treatment (%) | Outcome | |
|------------|------------|--------|-----------|---------------------------|-------------------------|----------------------|------------------------------|------------------------------|------------------|---------------------------|----------------------------|----------------------------|---------------------------|-----------------------------|-----------------------------|----------------------------|--------------|
| 0 | Patient ID | Age | Gender | Diagnosis | Symptom Severity (1-10) | Mood Score (1-10) | Sleep Quality (1-10) | Physical Activity (hrs/week) | Medication | Therapy Type | Treatment Start Date | Treatment Duration (weeks) | Stress Level (1-10) | Treatment Progress (1-10) | AI-Detected Emotional State | Adherence to Treatment (%) | Outcome |
| 1 | 1 | 43 | Female | Major Depressive Disorder | 10 | 5 | 8 | 5 | Mood Stabilizers | Interpersonal Therapy | 25/01/2024 | 11 | 9 | 7 | Anxious | 66 | Deteriorated |
| 2 | 2 | 40 | Female | Major Depressive Disorder | 9 | 5 | 4 | 7 | Antipsychotics | Interpersonal Therapy | 27/02/2024 | 11 | 7 | 7 | Neutral | 78 | No Change |
| 3 | 3 | 55 | Female | Major Depressive Disorder | 6 | 3 | 4 | 3 | SSRIs | Mindfulness-Based Therapy | 20/03/2024 | 14 | 7 | 5 | Happy | 62 | Deteriorated |
| 4 | 4 | 34 | Female | Major Depressive Disorder | 6 | 3 | 6 | 5 | SSRIs | Mindfulness-Based Therapy | 29/03/2024 | 8 | 8 | 10 | Excited | 72 | Deteriorated |

Gambar 3. 1 Data file CSV

Pada gambar 1 diatas, terdapat dataset terapi kesehatan mental, terdapat 500 data dengan 17 atribut. *Patient ID* merupakan id unik dari setiap pasien. *Age* merupakan usia pasien. *Gender* menunjukkan jenis kelamin pasien. *Diagnosis* merupakan jenis kondisi kesehatan mental yang didiagnosis pada pasien, seperti *Major Depressive Disorder* atau *Panic Disorder*. *Symptom Severity* (1-10) adalah tingkat keparahan gejala pasien berdasarkan skala 1-10, di mana nilai lebih tinggi menunjukkan keparahan yang lebih besar. *Mood Score* (1-10) adalah skor yang menggambarkan tingkat mood atau emosi pasien pada skala 1-10, di mana nilai lebih tinggi menunjukkan mood yang lebih baik. *Sleep Quality* (1-10) menunjukkan kualitas tidur pasien pada skala 1-10, di mana nilai lebih tinggi berarti kualitas tidur yang lebih baik. *Physical Activity* (*hrs/week*) menunjukkan jumlah rata-rata aktivitas fisik yang dilakukan pasien dalam jam per minggu. *Medication* adalah jenis obat yang dikonsumsi pasien, seperti *Antidepressants* atau *Mood Stabilizers*. *Therapy Type* menunjukkan jenis terapi yang diterima pasien, seperti *Interpersonal Therapy* atau *Cognitive Behavioral Therapy*. *Treatment Start Date* adalah tanggal dimulainya terapi. *Treatment Duration* (*weeks*) adalah durasi terapi dalam minggu. *Stress Level* (*1-10*) adalah tingkat stres pasien berdasarkan skala 1-10, di mana nilai lebih tinggi menunjukkan tingkat stres yang lebih besar. *Outcome* menunjukkan

hasil terapi, seperti *Improved*, *No Change*, atau *Deteriorated*. *Treatment Progress* (1-10) menggambarkan perkembangan pasien selama terapi pada skala 1-10. *AI-Detected Emotional State* adalah kondisi emosional pasien yang dianalisis oleh sistem berbasis AI, seperti *Anxious* atau *Happy*. *Adherence to Treatment* (%) adalah persentase tingkat kepatuhan pasien terhadap terapi yang dijalani.

3.2 Alur proses

Alur proses merupakan gambaran dari suatu proses yang akan dilakukan dalam sebuah pemecahan masalah, proses ini akan menjelaskan secara rinci mengenai tahapan dari penelitian. Dengan adanya alur proses yang jelas ini maka akan memudahkan pemahaman mengenai tahapan yang harus dilakukan dalam menganalisis data dan menghasilkan klasifikasi yang akurat mengenai hasil pengobatan.

3.2.1 Identifikasi Dan Analisis Proses Bisnis

Identifikasi dan analisis proses klasifikasi merupakan bagian penting dalam memahami cara kerja sistem prediksi hasil pengobatan kesehatan mental. Kedua konsep ini sering diterapkan secara berurutan untuk memberikan gambaran lengkap tentang bagaimana sistem menggunakan algoritma *Random Forest* dalam memprediksi hasil pengobatan berdasarkan data pasien.

a. Identifikasi proses bisnis

Identifikasi proses prediksi dilakukan dengan mengidentifikasi objek penelitian terkait, yaitu hasil pengobatan kesehatan mental pasien. Identifikasi proses prediksi akan menjabarkan langkah-langkah yang terkait dengan prediksi hasil

pengobatan menggunakan algoritma *Random Forest*. Untuk memprediksi hasil pengobatan kesehatan mental pasien, beberapa proses yang dilakukan antara lain:

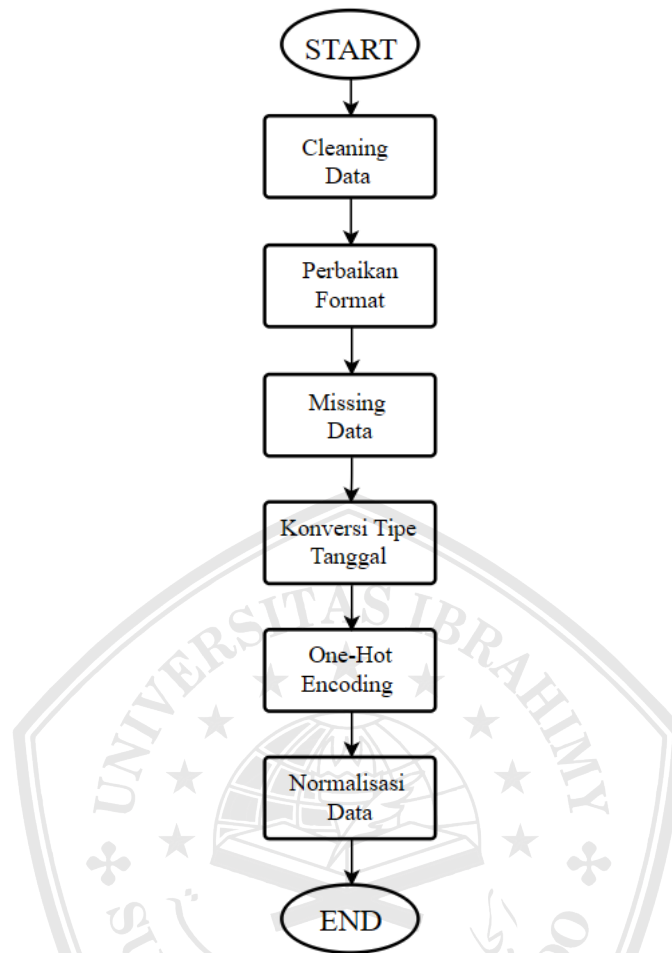
1. *Pre-Processing Data*
2. Pembagian Data (*Train-Test Split*)
3. Implementasi Algoritma *Random Forest*
4. Evaluasi Model
5. Hasil dan Solusi

3.2.2 Analisis Proses Bisnis

Analisis proses bisnis dalam penelitian ini dilakukan untuk memahami secara sistematis alur kerja dalam sistem pengelolaan pengobatan pasien dengan gangguan kesehatan mental, yang menjadi dasar dalam pengembangan model prediksi.

a. *Pre-processing Data*

Pre-processing data merupakan salah satu tahapan penting dalam proses analisis data, khususnya sebelum data digunakan dalam pemodelan algoritma. Tahapan ini bertujuan untuk membersihkan, merapikan, dan menyiapkan data agar layak untuk dianalisis. Hal ini dilakukan karena data mentah umumnya masih mengandung berbagai permasalahan, seperti data yang tidak lengkap, nilai yang menyimpang (*outlier*), serta format data yang belum konsisten [29].



Gambar 3. 2 Tahap *Pre-Processing*

1. *Cleaning Data* atau pembersihan data, yaitu dengan menghapus data duplikat, merapikan format kolom, serta memeriksa ketidaksesuaian data agar tidak memengaruhi hasil pemodelan.
2. Perbaikan Format, yaitu menyesuaikan kolom agar konsisten termasuk konversi tanggal ke format *datetime* dan perubahan tipe data sesuai kebutuhan.
3. *Missing Data*, yaitu mengidentifikasi dan menangani data yang kosong atau hilang (*missing values*). Penanganannya dapat berupa penghapusan baris data atau pengisian dengan nilai tertentu.

4. Konversi Tipe Data Tanggal, yaitu mengubah kolom *treatment start date* yang awalnya bertipe *string* menjadi format waktu
 5. *One-Hot Encoding*, proses ini digunakan untuk mengubah variable kategorikal menjadi representasi numerik biner agar dapat diproses oleh algoritma. Teknik yang digunakan adalah *One-Hot Encoding*, yang menghasilkan kolom baru untuk setiap kategori unik.
 6. Normalisasi Data, yaitu untuk meningkatkan performa model, fitur numerik distandarkan ke rentang yang sama.
- b. Pembagian Data (*Train-Test Split*)
- Setelah melalui tahap *pre-processing*, data yang telah bersih dan siap digunakan selanjutnya dibagi menjadi dua subset, yaitu data latih (*training data*) 80% dan data uji (*testing data*) 20%. Pembagian ini dilakukan untuk memisahkan data yang digunakan dalam proses pelatihan model dan data yang digunakan untuk menguji kinerja model, guna menghindari terjadinya *overfitting* serta mengevaluasi kemampuan generalisasi model terhadap data yang belum pernah dilatih sebelumnya.
- c. Implementasi algoritma *Random Forest*
- Adapun tahapan-tahapan implementasi algoritma *Random Forest* dalam penelitian ini adalah sebagai berikut :
1. Teknik *Bootstrap Sampling* untuk pemilihan subset data latih, kemudian dibentuk pohon keputusan untuk setiap subset dengan persamaan (2.1).
 2. Melakukan pemilihan fitur secara acak pada setiap node pohon keputusan dengan persamaan (2.2).

3. Menghitung probabilitas terbaik dari atribut dalam membagi data dengan persamaan (2.3).
4. Mengukur ketidakteraturan data dengan persamaan (2.4).
5. Menentukan hasil prediksi akhir berdasarkan voting mayoritas dari seluruh pohon untuk memutuskan diagnose akhir pasien dengan persamaan (2.5).

d. *Confusion Matrix*

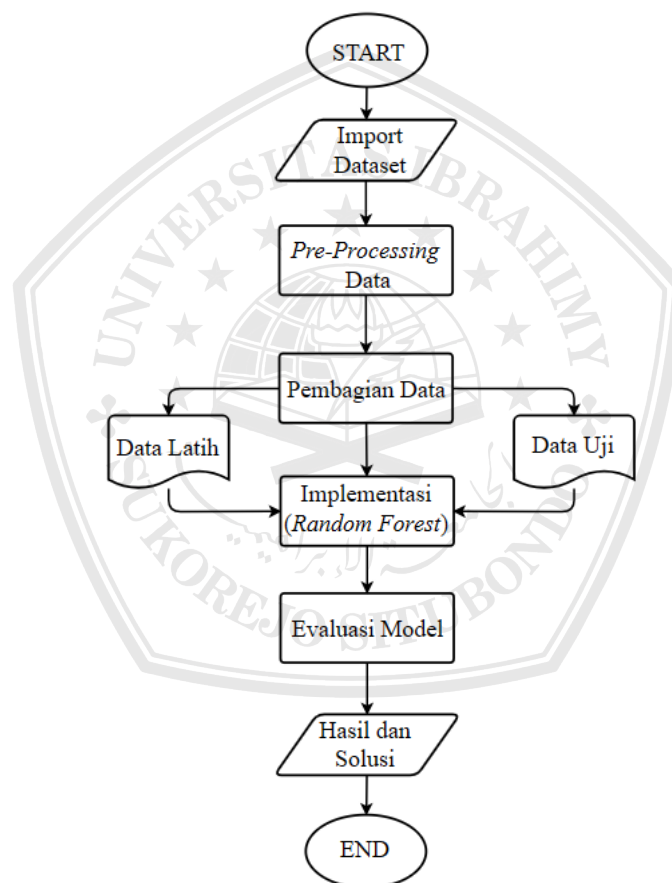
Model yang telah dibangun kemudian digunakan untuk melakukan prediksi, dan selanjutnya dievaluasi menggunakan sejumlah metrik untuk menilai kualitas hasil prediksi, yaitu :

1. Akurasi (*Accuracy*) bertujuan untuk menunjukkan persentase prediksi hasil pengobatan yang benar terhadap seluruh data, dengan persamaan (2.6)
2. Presisi (*Precision*) bertujuan untuk mengukur kemampuan model dalam menangkap seluruh kasus positif, dengan persamaan (2.7)
3. Recall (*Sensitivity*) bertujuan untuk mengukur kemampuan model dalam menangkap seluruh kasus positif, dengan persamaan (2.8)
4. F1-Score bertujuan untuk menyeimbangkan antara nilai Presisi dan Recall, terutama ketika distribusi kelas target tidak seimbang, dengan persamaan (2.9)

e. Flowchart Sistem

Pada *flowchart* sistem ini menggambarkan bahwa proses dimulai dengan menginputkan dataset diagnosis dan perawatan kesehatan mental pasien, kemudian data tersebut di proses melalui tahap pengolahan data seperti pembersihan data, normalisasi data dan pengkodean data. Selanjutnya data

dibagi menjadi dua bagian yaitu data latih dan data uji, Selanjutnya eksperimen dilakukan menggunakan algoritma *Random Forest* untuk mengidentifikasi diagnosis dan pengobatan kesehatan mental. Setelah eksperimen selesai dilakukan evaluasi model menggunakan *confusion matrix*, setelah semua proses selesai dilakukan maka dapat menentukan hasil dan solusi dari eksperimen yang dilakukan.



Gambar 3. 3 Flowchart Sistem

3.2.3 Identifikasi Dan Analisis Kebutuhan

a. Identifikasi Kebutuhan Fungsional

Sebuah proses akan berjalan lancar jika kebutuhan fungsional terpenuhi. Dimana dalam konteks penelitian ini, kebutuhan tersebut mencakup data diagnosis dan perawatan kesehatan mental serta fungsi-fungsi utama seperti *import* dataset pasien, *pre-processing* data, pembangunan model prediktif dengan algoritma *Random Forest*, prediksi hasil pengobatan, dan evaluasi performa model menggunakan metrik klasifikasi yang relevan.

b. Analisis Kebutuhan Fungsional

Analisis kebutuhan fungsional adalah menggambarkan fungsi-fungsi utama yang harus tersedia agar sistem dapat berjalan dengan baik. pada tahap ini data akan diolah menggunakan sebuah tahapan data yakni :

1. Import Dataset

Pada tahap pertama ini ialah mengimport dataset mentah kedalam sistem untuk diproses lebih lanjut oleh sistem.

2. *Pre-processing* Data

Tahapan berikutnya adalah *pre-processing* data, yang mencakup pembersihan data (*data cleaning*), penanganan nilai kosong (*missing values*), pengkodean variabel kategorikal, serta pembagian data menjadi data latih dan data uji. Sistem perlu menyediakan fungsi otomatis untuk menjalankan proses ini agar data siap digunakan oleh algoritma.

3. Pembagian Data

Tahap berikutnya adalah pembagian data yang mana data dibagi menjadi dua bagian yaitu data latih (*Training*) dan data uji (*Testing*). Pembagian ini dilakukan untuk memisahkan data yang digunakan dalam proses pelatihan model dan data yang digunakan untuk menguji kinerja model.

4. Pelatihan Model *Random Forest*

Setelah data dipersiapkan, sistem harus mampu melakukan proses pelatihan (*training*) terhadap data latih dan pengujian menggunakan algoritma *Random Forest*. Ini mencakup pembentukan beberapa pohon keputusan dan penggunaan metode voting untuk menghasilkan prediksi hasil pengobatan.

5. Evaluasi Model

Setelah model selesai dilatih, sistem harus mampu melakukan evaluasi model yang memungkinkan pengguna melihat performa model dengan menggunakan metrik klasifikasi seperti *akurasi*, *presisi*, *recall*, dan *F1-Score*.

6. Hasil dan Solusi

Tahapan selanjutnya, sistem diharapkan dapat menyajikan hasil prediksi dan analisis solusi yang dapat digunakan oleh pihak terkait (misalnya tenaga medis atau peneliti) untuk menilai apakah pengobatan pasien diprediksi akan berhasil atau tidak. Informasi ini dapat mendukung pengambilan keputusan yang lebih tepat dalam strategi penanganan kesehatan mental.

3.2.4 Identifikasi Dan Analisis Alternatif Solusi

a. Identifikasi Alternatif Solusi

Identifikasi alternatif solusi adalah proses mengevaluasi beberapa pendekatan yang dapat digunakan dalam menyelesaikan permasalahan yang telah diidentifikasi.

Tabel 3. 1 Identifikasi Alternatif Solusi

| Karakteristik | Alternatif 1 (Python) | Alternatif 2 (Rapid Miner) |
|---|---|---|
| Bagian sistem yang terkomputerisasi. | Seluruh proses dari import data, <i>pre-processing</i> , pelatihan model <i>Random Forest</i> , hingga evaluasi dilakukan secara otomatis melalui skrip python di <i>Goggle Colab</i> | Semua proses dilakukan melalui <i>workflow drag-and-drop</i> di antarmuka visual <i>RapidMiner Studio</i> . |
| Keuntungan, deskripsi ringkas keuntungan yang akan direalisasikan | Gratis, tidak perlu instalasi lokal, fleksibel untuk integrasi dan kolaborasi serta cocok untuk pemrograman lanjutan | Mudah digunakan, cocok untuk pengguna non-teknis serta cepat membangun model tanpa coding |
| Perangkat lunak aplikasi | <i>Goggle Colab</i> | <i>RapidMiner Studio</i> |
| Metode pemrosesan data | <i>Batch processing</i> dan <i>real-time</i> | <i>Batch</i> dan <i>real-time</i> secara visual melalui modul proses |
| Alat output | Monitor, printer dan disimpan di <i>goggle drive</i> | Monitor, printer dan dapat ditampilkan langsung di <i>rapidminer</i> |
| Alat input | <i>Keyboard, mouse</i> dan diakses dari <i>goggle drive</i> | <i>Keyboard, mouse</i> dan <i>Microsoft excel</i> |
| Alat penyimpanan | Data disimpan dalam <i>goggle drive</i> | Data disimpan dalam label <i>repository</i> |

b. Analisis Kelayakan Alternatif Solusi

Analisis kelayakan berfungsi untuk menilai sejauh mana alternatif solusi yang telah diidentifikasi dapat diimplementasikan secara efektif dalam penelitian ini, dengan menganalisis sistem yang akan dibangun berdasarkan dokumen dan data yang telah dikumpulkan sebelumnya.

Tabel 3. 2 Analisis Kelayakan

| Kriteria kelayakan | Bobot | Alternatif 1 (Python) | Alternatif 2 (RapidMiner) |
|---------------------------------------|-------|---|---|
| Kelayakan operasional fungsional | 40% | Memenuhi seluruh kebutuhan fungsional sistem dan dapat dikerjakan oleh banyak orang | Mudah digunakan tetapi terbatas pada <i>user interface</i> tertentu |
| Skor | | 38 | 30 |
| Kelayakan teknis teknologi | 30% | Teknologi open-source yang lengkap | Teknologi yang digunakan sudah cukup memadai |
| Skor | | 30 | 25 |
| Kelayakan ekonomis biaya pengembangan | 30% | Gratis, berbasis open-source, tidak memerlukan pengembangan tambahan, dan mendukung kolaborasi online | Bersifat komersial, versi gratis sangat terbatas |
| Skor | | 30 | 20 |
| Total | 100% | 98 | 75 |

3.3 Desain Proses

Desain proses merupakan tahapan dalam merancang sebuah sistem secara sistematis, agar setiap proses berjalan sesuai alur yang diinginkan dan saling terhubung. Dalam konteks penelitian ini, desain proses menggambarkan alur kerja dari pengolahan data hingga menghasilkan prediksi menggunakan algoritma

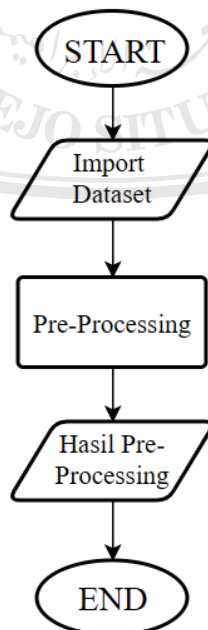
Random Forest untuk memprediksi hasil pengobatan kesehatan mental berdasarkan dataset yang tersedia.

a. Identifikasi Desain Proses

Identifikasi desain proses merupakan tahap awal dalam merancang alur kerja sistem yang akan dibangun, berdasarkan kebutuhan dan tujuan dari penelitian. Tahapan ini bertujuan untuk menggambarkan secara umum langkah-langkah yang dibutuhkan agar sistem dapat berjalan secara sistematis, efektif, dan sesuai dengan kebutuhan analisis prediktif. Berikut adalah desain proses yang dilakukan :

1. *Pre-Processing* data

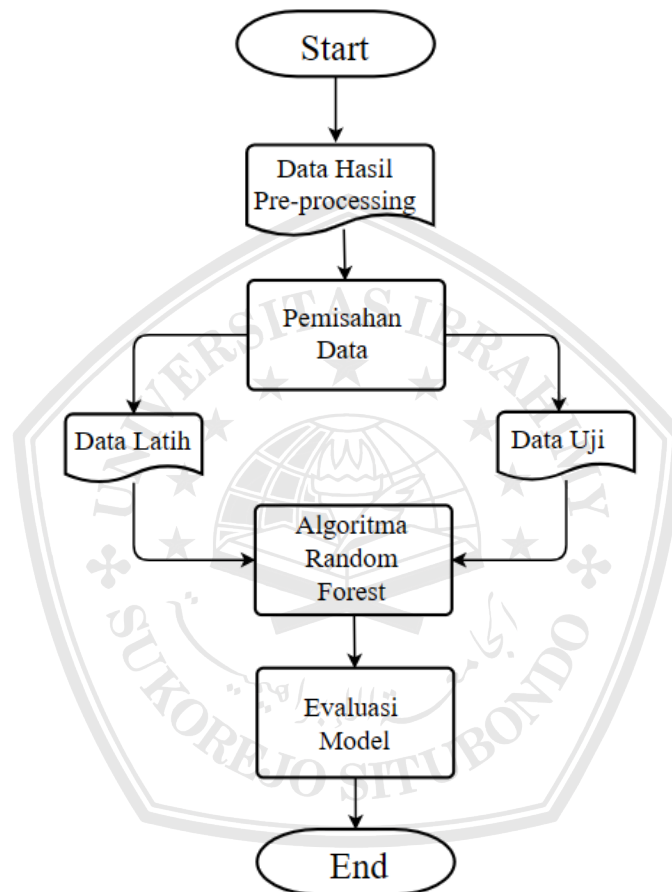
Pada proses *pre-processing* data ini dilakukan untuk mengolah dataset mentah yang diperoleh dari platform *Kaggle* yang berbentuk *excel* agar siap digunakan dalam perhitungan algoritma *Random Forest*.



Gambar 3. 4 *Pre-Processing* Data

2. Metode Random Forest

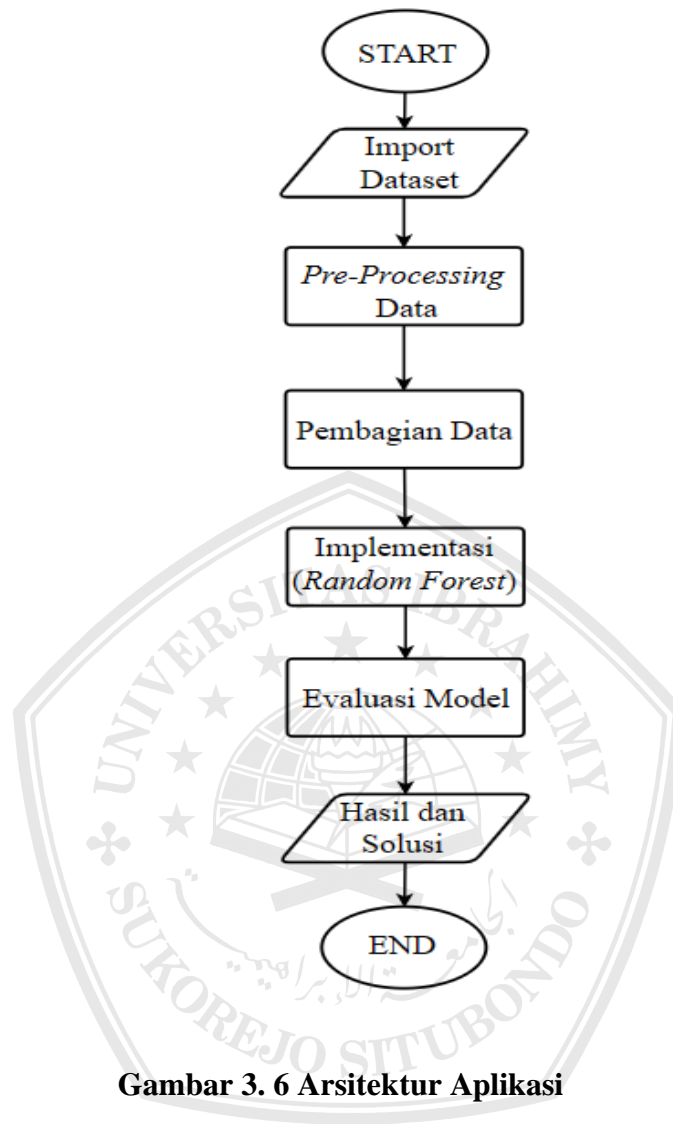
Pada proses ini bertujuan untuk menggambarkan secara sistematis alur penerapan algoritma *Random Forest* dalam membangun model prediksi keberhasilan pengobatan kesehatan mental.



Gambar 3. 5 Metode *Random Forest*

b. Arsitektur Aplikasi

Arsitektur sistem berfungsi untuk menjelaskan bentuk keseluruhan sistem prediksi berbasis algoritma *Random Forest* agar setiap proses dapat berjalan secara terstruktur dan mudah dipahami.



Gambar 3. 6 Arsitektur Aplikasi

BAB IV

IMPLEMENTASI SISTEM

4.1 Konstruksi Sistem

4.1.1 Kebutuhan Sistem

Kebutuhan sistem merupakan proses yang harus dipenuhi agar sistem dapat berjalan secara optimal dan mampu mendukung seluruh proses. Kebutuhan ini dibagi menjadi dua bagian, yaitu kebutuhan perangkat keras (*hardware*) dan perangkat lunak (*software*). Berikut adalah perangkat yang digunakan :

1. Hardware
 - a. Laptop HP
 - b. *Mouse dan Keyboard*
 - c. *Processor Intel Core i5*
 - d. RAM minimum 8 GB
 - e. SSD dengan kapasitas 256 GB
2. Software
 - a. Windows 10
 - b. *Google Colab*
 - c. *Microsoft Excel*
 - d. *Google Chrome*

4.2 Pengumpulan Data

Pada tahap awal dalam penelitian yakni data yang digunakan merupakan data sekunder yang bersumber dari *platform Kaggle*, sebuah situs yang menyediakan

berbagai dataset terbuka untuk keperluan riset dan pengembangan sistem berbasis data. *Dataset* yang digunakan adalah dataset *Mental Health Diagnosis And Treatment Monitoring* yang mempresentasikan kondisi nyata pasien kesehatan mental. dataset ini terdiri dari 500 data dengan 17 atribut yang diunduh dalam format CSV dan digunakan sebagai dasar untuk pelatihan dan pengujian model *Random Forest*. Dataset ini menjadi komponen penting dalam sistem yang akan diproses untuk menghasilkan prediksi keberhasilan pengobatan. Berikut tahap-tahap pengumpulan data :

1. Pencarian *dataset* dengan menggunakan kata kunci *mental health treatment monitoring*
2. Pengunduhan dataset *mental Health Diagnosis and Treatment Monitoring* dalam format CSV
3. Pemeriksaan struktur data untuk memastikan kelengkapan kolom dan kesesuaian dengan kebutuhan penelitian
4. Penyimpanan data secara lokal maupun di *Google Drive*
5. Unggah dan baca dataset ke *Google Colab*
6. Penamaan kolom dataset sesuai dengan struktur data

Berikut adalah gambaran dari tahap-tahap pengumpulan data :

| Patient ID | Age | Gender | Diagnosis | Symptom Severity (1-10) | Mood Score (1-10) | Sleep Quality (1-10) | Physical Activity (hrs/week) | Medication | Therapy Type | Treatment Start Date | Treatment Duration (weeks) | Stress Level (1-10) | Treatment Progress (1-10) | AI-Detected Emotional State | Adherence to Treatment (%) | Outcome |
|------------|-----|--------|---------------------------|-------------------------|-------------------|----------------------|------------------------------|------------------|--------------------------------|----------------------|----------------------------|---------------------|---------------------------|-----------------------------|----------------------------|--------------|
| 1 | 43 | Female | Major Depressive Disorder | 10 | 5 | 8 | 5 | Mood Stabilizers | Interpersonal Therapy | 25/01/2024 | 11 | 9 | 7 | Anxious | 66 | Deteriorated |
| 2 | 40 | Female | Major Depressive Disorder | 9 | 5 | 4 | 7 | Antipsychotics | Interpersonal Therapy | 27/02/2024 | 11 | 7 | 7 | Neutral | 78 | No Change |
| 3 | 55 | Female | Major Depressive Disorder | 6 | 3 | 4 | 3 | SSRIs | Mindfulness-Based Therapy | 20/03/2024 | 14 | 7 | 5 | Happy | 62 | Deteriorated |
| 4 | 34 | Female | Major Depressive Disorder | 6 | 3 | 6 | 5 | SSRIs | Mindfulness-Based Therapy | 29/03/2024 | 8 | 8 | 10 | Excited | 72 | Deteriorated |
| 5 | 52 | Male | Panic Disorder | 7 | 6 | 6 | 8 | Anxiolytics | Interpersonal Therapy | 18/03/2024 | 12 | 5 | 6 | Excited | 63 | Deteriorated |
| 6 | 28 | Male | Panic Disorder | 8 | 7 | 6 | 4 | SSRIs | Cognitive Behavioral Therapy | 11/01/2024 | 13 | 9 | 7 | Stressed | 82 | No Change |
| 7 | 59 | Male | Generalized Anxiety | 6 | 6 | 5 | 3 | Mood Stabilizers | Dialectical Behavioral Therapy | 21/02/2024 | 13 | 8 | 5 | Anxious | 79 | Deteriorated |
| 8 | 32 | Female | Major Depressive Disorder | 5 | 3 | 5 | 2 | Antidepressants | Cognitive Behavioral Therapy | 03/02/2024 | 12 | 5 | 6 | Stressed | 74 | No Change |
| 9 | 21 | Female | Major Depressive Disorder | 6 | 7 | 7 | 10 | Antidepressants | Mindfulness-Based Therapy | 16/01/2024 | 15 | 5 | 6 | Neutral | 67 | No Change |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 500 | 58 | Male | Bipolar Disorder | 5 | 7 | 6 | 2 | SSRIs | Cognitive Behavioral Therapy | 24/03/2024 | 10 | 6 | 5 | Happy | 70 | Deteriorated |

Gambar 4. 1 Dataset Mental Health Diagnosis And Treatment Monitoring

Segmen 1. Unggah dan Baca Dataset

```
from google.colab import files
uploaded = files.upload()

import pandas as pd
import io

df = pd.read_csv(io.BytesIO(uploaded['klasifikasi (3).csv']), delimiter=';', header=None)
```

Segmen 2. Penamaan kolom

```
columns = ['Patient ID', 'Age', 'Gender', 'Diagnosis', 'Symptom Severity (1-10)', 'Mood Score (1-10)', 'Sleep Quality (1-10)', 'Physical Activity (hrs/week)', 'Medication', 'Therapy Type', 'Treatment Start Date', 'Treatment Duration (weeks)', 'Stress Level (1-10)', 'Treatment Progress (1-10)', 'AI-Detected Emotional State', 'Adherence to Treatment (%)', 'Outcome']

df.columns = columns

df.head()
```

4.3 Pre-processing Data

Pre-processing data merupakan proses menyiapkan data agar dapat digunakan oleh algoritma secara optimal. Pada penelitian ini, *Pre-processing* dilakukan melalui beberapa tahap berikut :

1. Cleaning

Pada tahap awal ini dilakukan *Cleaning* data atau pembersihan data yang bertujuan untuk memastikan bahwa dataset yang digunakan bersih, konsisten, bebas dari kesalahan yang dapat mengganggu proses analisis. Dalam tahap ini dilakukan menghapus baris header duplikat, memberi nama kolom, serta memastikan format data sesuai. Nilai kosong pada kolom numerik diisi dengan rata-rata, sedangkan pada kolom kategorikal diisi dengan nilai modus. Kolom tanggal juga dikonversi ke format *datetime*. Hasil *cleaning* menghasilkan data yang bersih dan siap diproses lebih lanjut. Berikut adalah kode *Cleaning* data

Segmen 3. *Cleaning* Data

```
header_values = df.iloc[0].values
df_cleaned = df[~df.eq(header_values).all(axis=1)].copy()

df_cleaned.reset_index(drop=True, inplace=True)

df_cleaned.columns = ['Patient ID', 'Age', 'Gender',
'Diagnosis', 'Symptom Severity (1-10)', 'Mood Score (1-10)',
'Sleep Quality (1-10)', 'Physical Activity
(hrs/week)', 'Medication', 'Therapy Type', 'Treatment Start
Date', 'Treatment Duration (weeks)', 'Stress Level (1-
10)', 'Treatment Progress (1-10)', 'AI-Detected Emotional
State', 'Adherence to Treatment (%)', 'Outcome'
]
```

Segmen Lanjutan 3. *Cleaning Data*

```
numerical_cols = [  
    'Age', 'Symptom Severity (1-10)', 'Mood Score (1-10)',  
    'Sleep Quality (1-10)', 'Physical Activity (hrs/week)',  
    'Treatment Duration (weeks)', 'Stress Level (1-10)',  
    'Treatment Progress (1-10)', 'Adherence to Treatment (%)']  
  
for col in numerical_cols:  
  
df_cleaned[col]=pd.to_numeric(df_cleaned[col],errors='coerce')  
  
df_cleaned[numerical_cols]=df_cleaned[numerical_cols].fillna(df_cleaned[numerical_cols].mean())  
  
categorical_cols = ['Gender', 'Diagnosis', 'Medication',  
    'Therapy Type', 'Outcome', 'AI-Detected Emotional State']  
  
for col in categorical_cols:  
  
    mode_value = df_cleaned[col].mode().iloc[0] if not df_cleaned[col].mode().empty else ''  
  
    df_cleaned[col] = df_cleaned[col].fillna(mode_value)  
  
df_cleaned['Treatment Start Date'] = pd.to_datetime(df_cleaned['Treatment Start Date'], errors='coerce')  
  
df_cleaned.info()
```

2. Perbaiki Format

Setelah proses *cleaning* dilakukan, langkah selanjutnya adalah melakukan perbaikan format data agar seluruh atribut pada dataset memiliki tipe data yang sesuai dan dapat digunakan dalam proses analisis lebih lanjut. Perbaikan format ini bertujuan untuk menyelaraskan struktur dataset dengan kebutuhan algoritma *machine learning*.

Segmen 4. Perbaiki Format

```
columns =['Patient ID', 'Age', 'Gender', 'Diagnosis',  
    'Symptom Severity (1-10)', 'Mood Score (1-10)', 'Sleep  
Quality (1-10)', 'Physical Activity  
(hrs/week)', 'Medication', 'Therapy Type', 'Treatment Start  
Date', 'Treatment Duration (weeks)', 'Stress Level (1-  
10)', 'Treatment Progress (1-10)', 'AI-Detected Emotional  
State', 'Adherence to Treatment (%)', 'Outcome']
```

Segmen Lanjutan 4. Perbaiki Format

```
numerical_cols = [  
    'Age', 'Symptom Severity (1-10)', 'Mood Score (1-  
10)', 'Sleep Quality (1-10)', 'Physical Activity  
(hrs/week)', 'Treatment Duration (weeks)', 'Stress Level (1-  
10)', 'Treatment Progress (1-10)', 'Adherence to Treatment  
(%)'  
]  
  
for col in numerical_cols:  
    df_cleaned[col] = pd.to_numeric(df_cleaned[col],  
errors='coerce')
```

3. Missing Data

Penanganan *Missing* data adalah salah satu tahapan penting dalam proses pengolahan data agar analisis yang dilakukan tidak terganggu oleh nilai-nilai kosong yang dapat menyebabkan kesalahan dalam perhitungan atau pelatihan model. Dalam hal ini ditemukan adanya nilai kosong pada beberapa kolom, baik numerik ataupun kategorikal. Untuk mengatasi hal tersebut, nilai kosong pada kolom numerik diisi dengan nilai rata-rata, sedangkan pada kolom kategorikal diisi dengan nilai modus. Metode ini dipilih agar tidak ada data yang dihapus dan informasi tetap utuh, sehingga dataset siap digunakan pada tahap selanjutnya.

Segmen 5. Missing Data

```
numeric_cols=df.select_dtypes(include=['number']).columns  
df[numeric_cols]=df[numeric_cols].fillna(df[numeric_cols].m  
ean())  
  
categorical_cols=df.select_dtypes(include=['object']).colum  
ns  
  
df[categorical_cols]=df[categorical_cols].fillna(df[categor  
ical_cols].mode().iloc[0])
```

4. Konversi Tipe Data Tanggal

Setelah penanganan *missing* data langkah selanjutnya adalah mengubah kolom *Treatment Start Date* yang awalnya bertipe teks (*string*) menjadi format waktu yang dikenali *Python* (*datetime*). proses konversi dilakukan menggunakan fungsi `pd.to_datetime()` agar dapat diolah lebih lanjut. Setelah dikonversi, informasi tanggal dipecah menjadi tiga kolom terpisah yakni *Treatment Start Day*, *Treatment Start Month*, Dan *Treatment Start Year*. Kolom tanggal asli kemudian dihapus karena sudah dipresentasikan dalam bentuk yang lebih terstruktur.

Segmen 6. Konversi Tipe Data Tanggal

```
df['Treatment Start Date']=pd.to_datetime(df['Treatment  
Start Date'], errors='coerce')  
  
df['Treatment Start Day'] = df['Treatment Start  
Date'].dt.day  
  
df['Treatment Start Month'] = df['Treatment Start  
Date'].dt.month  
  
df['Treatment Start Year'] = df['Treatment Start  
Date'].dt.year  
  
df.drop('Treatment Start Date', axis=1, inplace=True)
```

5. One-Hot Encoding

Pada tahap selanjutnya dilakukan *One-Hot Encoding* yakni untuk mengubah data kategorikal menjadi bentuk numerik agar dapat diproses oleh algoritma *machine learning*. Dalam setiap kategori unik dalam kolom akan diwakili oleh satu kolom baru yang bernilai 1 jika cocok, dan 0 jika tidak. Dengan penerapan teknik ini, data kategorikal berhasil direpresentasikan secara numerik tanpa kehilangan informasi, dan siap digunakan dalam proses pelatihan model.

Segmen 7. One-Hot Encoding

```
df_encoded = pd.get_dummies(df, columns=[
    'Gender', 'Diagnosis', 'Medication', 'Therapy Type',
    'Outcome', 'AI-Detected Emotional State'
])

df_encoded.head()
```

6. Normalisasi Data

Pada tahap normalisasi ini dilakukan untuk menyamakan skala antar fitur numerik agar tidak terjadi dominasi oleh fitur dengan nilai lebih besar. dalam hal ini, proses normalisasi dilakukan menggunakan Teknik standarisasi dengan bantuan fungsi *StandardScaler*. Fitur numerik yang telah dipilih kemudian diubah skalanya agar memiliki nilai rata-rata 0 dan standar deviasi 1. Langkah ini bertujuan untuk meningkatkan performa model dan mempercepat proses pelatihan.

Segmen 8. Normalisasi Data

```
from sklearn.preprocessing import StandardScaler

numerical_features =
df_encoded.select_dtypes(include=['number']).columns

scaler = StandardScaler()

df_encoded[numerical_features] =
scaler.fit_transform(df_encoded[numerical_features])

df_encoded.head()
```

4.4 Pembagian Data (*Train-Test Split*)

Setelah seluruh proses *pre-processing* selesai dilakukan, tahap selanjutnya adalah pembagian data yang dimana data dibagi menjadi dua bagian, yaitu data latih

(*Training data*) dan data uji (*Testing data*). Data latih digunakan untuk membangun dan melatih model *machine learning* agar dapat mengenali pola dan hubungan antar variable. Sedangkan, Data uji digunakan untuk mengukur performa model yang belum pernah dilatih sebelumnya. Pembagian ini dilakukan untuk memastikan bahwa model yang dibangun dapat dievaluasi secara objektif terhadap data yang belum pernah dilatih sebelumnya. Dalam hal ini, data akan dibagi dengan proporsi 80% untuk data latih dan 20% untuk data uji. Berikut adalah code dalam pembagian data :

Segmen 9. Pembagian Data

```
target_column = ['Outcome_Improved', 'Outcome_Deteriorated',  
'Outcome_No Change']  
X = df.drop(target_column, axis=1)  
y = df[target_column]  
  
X_train, X_test, y_train, y_test = train_test_split(  
    X, y, test_size=0.2, random_state=42)  
  
print("\nJumlah data training:", len(X_train))  
print("Jumlah data testing :", len(X_test))  
  
X_train.to_csv("X_train.csv", index=False)  
X_test.to_csv("X_test.csv", index=False)  
y_train.to_csv("y_train.csv", index=False)  
y_test.to_csv("y_test.csv", index=False)  
  
files.download("X_train.csv")  
files.download("X_test.csv")  
files.download("y_train.csv")  
files.download("y_test.csv")
```

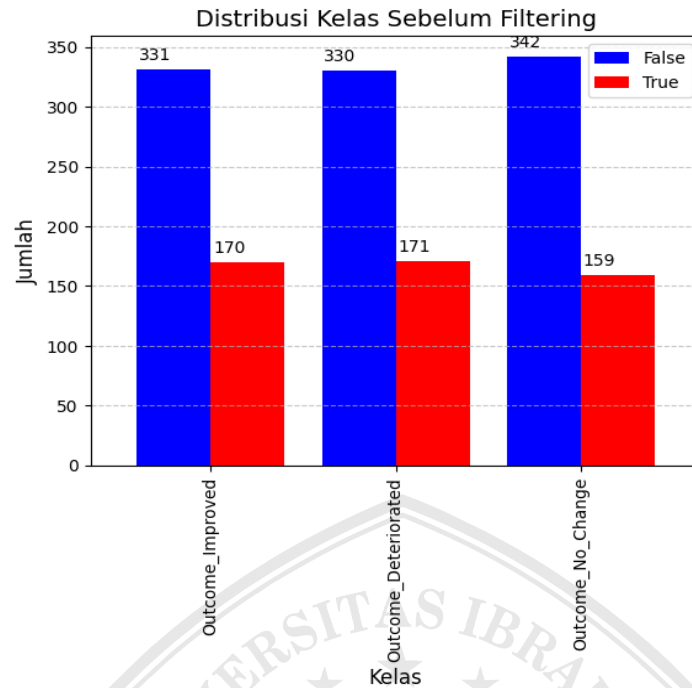
a. Hasil Distribusi Kelas Sebelum Filtering

Sebelum dilakukan proses *filtering* dan *preprocessing*, data terlebih dahulu dianalisis untuk mengetahui distribusi awal kelas pada label target. Tabel berikut menunjukkan pembagian data berdasarkan tiga kategori utama yaitu *Outcome_Improved*, *Outcome_Deteriorated*, dan *Outcome_No Change*. Masing-masing kelas terdiri dari data dengan *label True* dan *False*. Total keseluruhan data untuk setiap kelas adalah sebanyak 501 data, ditunjukkan pada tabel 4.1 berikut :

Tabel 4. 1 Pembagian Data Sebelum *Filtering*

| Label Kelas | Jumlah False | Jumlah True | Total |
|----------------------|--------------|-------------|-------|
| Outcome_Improved | 331 | 170 | 501 |
| Outcome_Deteriorated | 330 | 171 | 501 |
| Outcome_No Change | 342 | 159 | 501 |

Berdasarkan tabel di atas, dapat dilihat bahwa jumlah data pada masing-masing kelas relatif seimbang, baik dari segi total maupun perbandingan antara *label True* dan *False*. Informasi ini menjadi penting untuk memastikan bahwa tidak terjadi ketimpangan signifikan dalam distribusi kelas yang dapat memengaruhi performa model pada tahap pelatihan. Berikut adalah gambar distribusi kelas sebelum *filtering* dalam bentuk diagram :



Gambar 4. 2 Distribusi Kelas Sebelum Filtering

b. Hasil Distribusi Jumlah Data Training Dan Testing

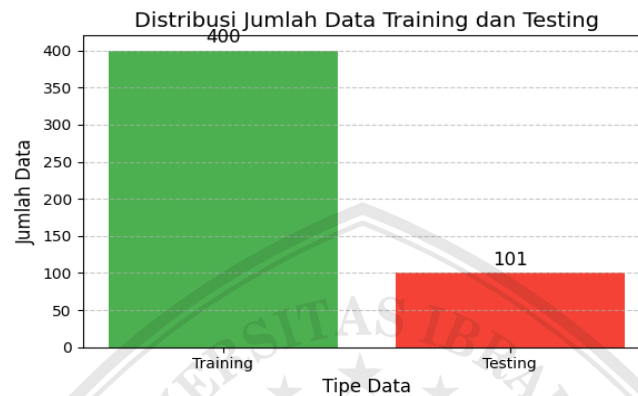
Proses pembagian data dilakukan dengan metode *stratifikasi* agar proporsi kelas pada data *training* dan data *testing* tetap seimbang. Teknik *stratifikasi* ini penting untuk menjaga distribusi kelas yang merata, terutama pada *dataset* yang memiliki beberapa label target. Berdasarkan hasil pembagian, diperoleh 400 data untuk keperluan pelatihan (*training*) dan 101 data untuk pengujian (*testing*). Rincian hasil pembagian data dapat dilihat pada Tabel berikut:

Tabel 4. 2 Pembagian Data Setelah Startifikasi

| Jenis Data | Jumlah Data |
|---------------|-------------|
| Data Training | 400 |
| Data Testing | 101 |
| Total | 501 |

Dengan pembagian ini, model dapat dilatih menggunakan data yang cukup *representatif*, sekaligus diuji menggunakan data yang belum pernah dilihat sebelumnya, sehingga evaluasi performa model menjadi lebih objektif dan valid.

Berikut adalah gambar diagram distribusi data *training* dan *testing* :



Gambar 4. 3 Split Data

4.5 Metode Random Forest

Setelah seluruh proses *pre-processing* dan pembagian data selesai dilakukan, tahap selanjutnya adalah implementasi algoritma *Random Forest* untuk membangun model klasifikasi hasil pengobatan kesehatan mental, yang mana metode ini dipilih karena kemampuannya dalam menangani dataset kompleks dengan banyak fitur serta memberikan akurasi tinggi melalui proses voting dari sejumlah pohon keputusan, dan model dilatih menggunakan 400 data latih dari total 501 data yang tersedia dengan menerapkan teknik *bootstrap sampling* (2.1) untuk membentuk beberapa subset acak yang digunakan dalam pembangunan masing-masing pohon, di mana pada setiap node hanya sebagian fitur yang dipilih secara acak (2.2), dan pemisahan terbaik ditentukan menggunakan kriteria *Gini Impurity* (2.3) dan *Information Gain* (2.4).

Parameter model :

- a. Random ForestClassifier, digunakan untuk membuat model *Random Forest*
- b. n_estimators=100, menunjukkan 100 pohon keputusan yang akan dilatih
- c. random_state=42, untuk memastikan hasil konsisten setiap kali model dijalankan
- d. fit(), untuk melatih model pada data latih (x_train dan y_train)
- e. predict(), untuk menghasilkan prediksi berdasarkan data uji(x_test), yang hasilnya disimpan dalam y_pred

berikut adalah kode program untuk membangun model *Random Forest* :

Segmen 10. Model Random Forest

```
rf_model = RandomForestClassifier(n_estimators=100,  
random_state=42)  
rf_model.fit(X_train, y_train)  
  
y_pred = rf_model.predict(X_test)
```

Setelah model *Random Forest* dilatih, proses prediksi dilakukan terhadap data uji untuk mengukur kemampuan model dalam memprediksi hasil pengobatan kesehatan mental. Hasil prediksi kemudian dievaluasi menggunakan beberapa metrik penting, yaitu : akurasi, precision, recall, f1-score serta analisis entropi dan information gain.

- a. Akurasi

Akurasi dalam penelitian ini digunakan digunakan untuk mengukur sejauh mana model *Random Forest* mampu memprediksi hasil pengobatan kesehatan mental secara tepat, baik untuk pasien yang mengalami perbaikan maupun yang tidak.

Segmen 11. Akurasi

```
from sklearn.metrics import accuracy_score
accuracy = accuracy_score(y_test, y_pred)
print(f"Akurasi Model: {accuracy * 100:.2f}%")
```

Hasil :

```
Akurasi Model: 99.01%
```

b. Precision, recall dan f1-score

Precision mengukur ketepatan prediksi positif, *Recall* mengukur seberapa banyak kasus positif yang berhasil dikenali model, sedangkan *F1-Score* merupakan rata-rata harmonis dari keduanya. Ketiga metrik ini memberikan gambaran menyeluruh tentang kinerja model dalam memprediksi hasil pengobatan kesehatan mental.

Segmen 12. Precision, Recall dan F1-Score

```
from sklearn.metrics import accuracy_score,
classification_report

accuracy = accuracy_score(y_test, y_pred)
print(f"Akurasi Model: {accuracy:.4f}\n")

target_names = ['Improved', 'Deteriorated', 'NoChange']
print("Laporan Klasifikasi:")
print(classification_report(
    y_test, y_pred,
    target_names=['Improved', 'Deteriorated', 'NoChange'],
    labels=[0, 1, 2] # sesuaikan dengan label numerik di
    y_test
))
```

Segmen Lanjutan 12. Precision, Recall dan F1-Score

| | | | | |
|-----------------------|-----------|--------|----------|---------|
| Akurasi Model: 0.9901 | | | | |
| Laporan Klasifikasi: | | | | |
| | precision | recall | f1-score | support |
| Improved | 1.00 | 0.98 | 0.99 | 65 |
| Deteriorated | 0.97 | 1.00 | 0.99 | 36 |
| NoChange | 0.00 | 0.00 | 0.00 | 0 |
| accuracy | | | 0.99 | 101 |
| macro avg | 0.66 | 0.66 | 0.66 | 101 |
| weighted avg | 0.99 | 0.99 | 0.99 | 101 |

a. Precision

Pada hasil *precision* diatas menunjukkan bahwa tingkat ketepatan model dalam memprediksi kelas positif sangat tepat dalam mengklasifikasi hasil pengobatan pasien. Pada kelas *Improved* memiliki precision sebesar 1.00, Pada kelas *Deteriorated* menunjukkan hasil 0.97 yang berarti prediksi model sangat akurat.

b. Recall

Recall mengukur kemampuan model dalam mengenali semua data yang benar-benar termasuk ke dalam suatu kelas. Pada hasil ini, *recall* untuk kelas *Improved* adalah 0.98 dan untuk *Deteriorated* adalah 1.00, yang berarti model hampir selalu berhasil mendeteksi pasien yang mengalami perbaikan dan sepenuhnya mendeteksi pasien yang memburuk. Nilai *recall* yang tinggi menunjukkan bahwa model sangat baik dalam menangkap seluruh kasus yang relevan.

c. F1-score

F1-score adalah metrik yang menggabungkan precision dan recall dalam satu nilai rata-rata harmonis. Pada hasil ini, *f1-score* untuk kelas *Improved* adalah 0.99 dan untuk *Deteriorated* juga 0.99, yang menunjukkan keseimbangan antara ketepatan dan kemampuan model dalam mengenali kasus sebenarnya. Nilai *F1-score* yang tinggi menandakan bahwa model memiliki performa yang sangat baik dan konsisten dalam memprediksi hasil pengobatan kesehatan mental.

c. Confusion matrix

Confusion matrix digunakan untuk menampilkan performa model klasifikasi dalam bentuk tabel, yang menunjukkan jumlah prediksi benar dan salah pada setiap kelas.

Segmen 13. Confusion Matrix

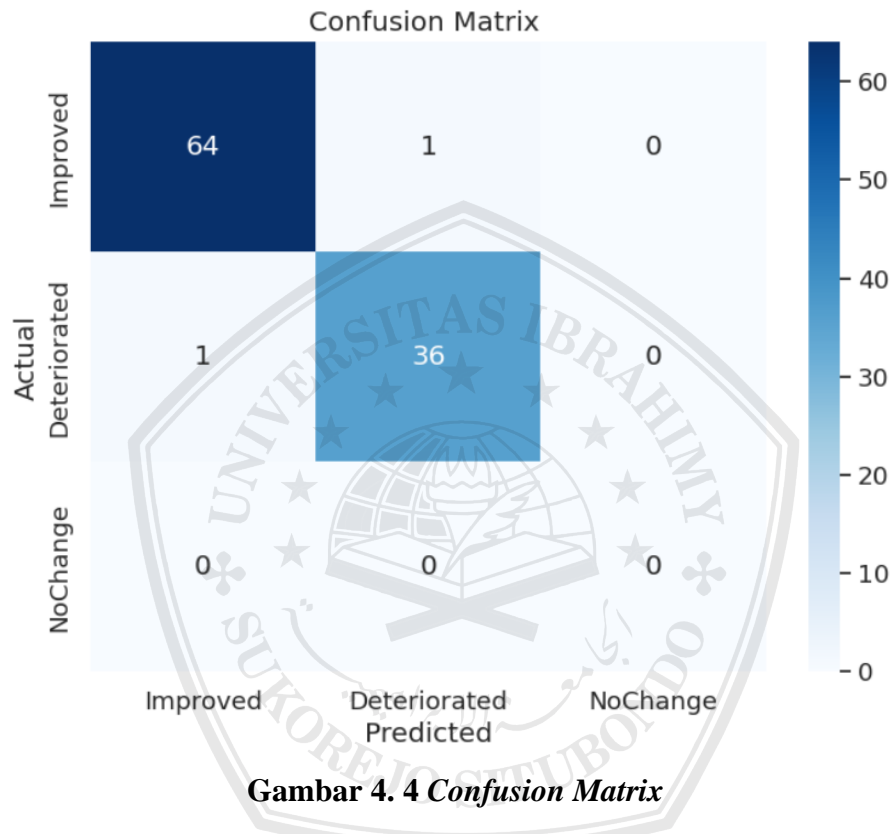
```
import numpy as np

cm_full = np.array([[64, 1, 0], # Improved
                   [1, 36, 0], # Deteriorated
                   [0, 0, 0]   # NoChange])

labels_full = ['Improved', 'Deteriorated', 'NoChange']
plt.figure(figsize=(8, 6))
sns.heatmap(cm_full, annot=True, fmt='d', cmap='Blues',
            xticklabels=labels_full,
            yticklabels=labels_full)
plt.title('Confusion Matrix ')
plt.xlabel('Predicted')
plt.ylabel('Actual')
plt.show()
```

Segmen Lanjutan 13. Confusion Matrix

```
df_cm_full = pd.DataFrame(cm_full, index=labels_full,  
                           columns=labels_full)  
  
print("Confusion Matrix :")  
  
print(df_cm_full)
```

**Gambar 4. 4 Confusion Matrix**

Confusion matrix diatas menunjukkan hasil prediksi model terhadap tiga kelas: *Improved* (membaik), *Deteriorated* (memburuk), dan *No Change* (tidak ada perubahan). Baris mewakili label sebenarnya (*true labels*), sedangkan kolom mewakili label yang diprediksi oleh model (*predicted labels*). Berikut penjelasan untuk setiap baris dan kolom :

a. Improved (membaik)

- 1) Jumlah data aktual sebanyak 65 data untuk kelas *Improved*
- 2) 64 data berhasil diprediksi dengan benar oleh model sebagai *Improved*

3) 1 data salah diklasifikasikan sebagai memburuk

b. Deteriorated (memburuk)

1) Terdapat 37 data aktual dengan label *Deteriorated*

2) 36 data berhasil diprediksi dengan benar sebagai *Deteriorated*

3) 1 data salah diklasifikasikan sebagai *Improved*

c. No Change (tidak ada perubahan)

1) 0 data aktual berlabel *No Change*

2) 0 data diprediksi dengan benar sebagai *No Change*

3) Tidak ada data yang salah diklasifikasikan sebagai *No Change*

d. Analisis Entropi Dan Information Gain

Dalam algoritma *Random Forest*, proses pemilihan fitur pada setiap node pohon didasarkan pada seberapa besar fitur tersebut dapat mengurangi ketidakpastian dalam data. Pengukuran ketidakpastian ini dikenal sebagai *Entropi*, sedangkan pengurangannya disebut *Information Gain*. Fitur dengan nilai *Information Gain* tertinggi akan dipilih sebagai pemisah pada node karena dianggap paling informatif. *Entropi* yang rendah menunjukkan data lebih seragam, sementara *Entropi* tinggi menunjukkan keragaman atau ketidakteraturan yang tinggi.

Dalam penelitian ini, konsep *Entropi* dan *Information Gain* menjadi dasar pembentukan struktur pohon dalam *Random Forest* untuk memprediksi hasil pengobatan kesehatan mental. Meskipun secara eksplisit tidak dihitung, nilai kontribusi fitur dapat dilihat melalui atribut `.feature_importances_` dalam pustaka *Scikit-learn*, yang mencerminkan seberapa besar pengaruh fitur tersebut dalam

mengurangi impurity di seluruh model. Berikut adalah fitur penting untuk *Random*

Forest :

Segmen 14. Kode Dan Hasil Entropi Dan Information Gain

```
importances = rf_model.feature_importances_  
  
feature_names = X.columns  
  
print(" abc Feature Importances:")  
for name, importance in zip(feature_names, importances):  
    print(f'{name}: {importance:.4f}')
```

```
Hasil Entropi dan Information Gain  
 abc Feature Importances:  
Age: 0.0432  
Symptom Severity (1-10): 0.0202  
Mood Score (1-10): 0.0251  
Sleep Quality (1-10): 0.0266  
Physical Activity (hrs/week): 0.0310  
Treatment Duration (weeks): 0.0269  
Stress Level (1-10): 0.0272  
Treatment Progress (1-10): 0.0259  
Adherence to Treatment (%): 0.0485  
Treatment Start Day: 0.0339  
Treatment Start Month: 0.0285  
Treatment Start Year: 0.0000  
Gender_Female: 0.0070  
Gender_Gender: 0.0000  
Gender_Male: 0.0085  
Diagnosis_Bipolar Disorder: 0.0074  
Diagnosis_Diagnosis: 0.0000  
Diagnosis_Generalized Anxiety: 0.0102
```

Segmen Lanjutan 14. Kode Dan Hasil Entropi Dan Information Gain

```
Diagnosis_Major Depressive Disorder: 0.0092
Diagnosis_Panic Disorder: 0.0069
Medication_Antidepressants: 0.0068
Medication_Antipsychotics: 0.0051
Medication_Anxiolytics: 0.0052
Medication_Benzodiazepines: 0.0071
Medication_Medication: 0.0000
Medication_Mood Stabilizers: 0.0079
Medication_SSRI: 0.0067
Therapy Type_Cognitive Behavioral Therapy: 0.0070
Therapy Type_Dialectical Behavioral Therapy: 0.0071
Therapy Type_Interpersonal Therapy: 0.0067
Therapy Type_Mindfulness-Based Therapy: 0.0075
Therapy Type_Therapy Type: 0.0000
AI-Detected Emotional State_AI-Detected Emotional State:
0.0000
AI-Detected Emotional State_Anxious: 0.0087
AI-Detected Emotional State_Depressed: 0.0044
AI-Detected Emotional State_Excited: 0.0085
AI-Detected Emotional State_Happy: 0.0054
AI-Detected Emotional State_Neutral: 0.0063
AI-Detected Emotional State_Stressed: 0.0076
Outcome_Deteriorated: 0.2775
Outcome_No Change: 0.2282
Outcome_Outcome: 0.0000
```

Feature Importance adalah metrik yang digunakan oleh algoritma *Random Forest* untuk mengukur kontribusi relatif setiap fitur terhadap akurasi prediksi model. Nilai yang lebih tinggi menunjukkan bahwa fitur tersebut lebih

berpengaruh dalam membantu model membuat keputusan klasifikasi. Berikut adalah uraian rinci berdasarkan kelompok fitur :

- a. *Age* : 0.053, usia pasien memberikan kontribusi sebesar 5.33%. dalam konteks ini menunjukkan umur menjadi faktor penting, meskipun tidak dominan, dalam mempengaruhi hasil prediksi pengobatan.
- b. *Adherence To Treatment (%)* : 0.0457, tingkat kepatuhan pasien terhadap pengobatan memberikan kontribusi sebesar 4.57%, yang mengindikasikan bahwa semakin tinggi kepatuhan, semakin besar pengaruhnya terhadap keberhasilan terapi.
- c. *Treatment Start Day* : 0.0406, Hari dimulainya pengobatan berkontribusi 4.06%, yang mengindikasikan bahwa waktu memulai terapi memiliki peran dalam hasil akhir pengobatan.
- d. *Treatment Duration (Week)* : 0.032, Durasi pengobatan menyumbang 3.21%, yang berarti lamanya waktu terapi juga diperhitungkan dalam model.
- e. Fitur-fitur ini masing-masing memberikan kontribusi antara 2.65%–3.11%, menunjukkan bahwa kondisi fisik dan emosional pasien turut memengaruhi klasifikasi meskipun tidak dominan.

1. *Physical Activity (Hrs/Week)* : 0.0311

2. *Stress Level (1-10)* : 0.0289

3. *Sleep Quality (1-10)* : 0.0287

4. *Mood Score (1-10)* : 0.0269

5. *Symptom Severity (1-10)* : 0.0265

- f. *AI-Detected Emotional State_Excited* : 0.0103, Ekspresi emosional pasien yang terdeteksi sebagai "*Excited*" oleh AI memberikan kontribusi 1.03%, yang berarti cukup membantu model dalam memahami kondisi psikologis pasien.
- g. Fitur-fitur lainnya seperti jenis kelamin, jenis obat, diagnosis, dan jenis terapi memiliki nilai antara 0.0048–0.0094, yang berarti berperan kecil dalam keputusan model.
1. *Diagnosis_Generalized Anxiety*: 0.0094
 2. *Medication_Mood Stabilizers*: 0.0090
 3. *Therapy Type_Mindfulness-Based Therapy*: 0.0089
 4. *Gender_Male*: 0.0082
- h. *Outcome_Deteriorated* : 0.2511, Fitur ini memberikan kontribusi sebesar 25.11% terhadap keputusan klasifikasi model. Ini menunjukkan bahwa riwayat pasien yang mengalami penurunan kondisi sangat berpengaruh dalam menentukan prediksi hasil pengobatan.
- i. *Outcome_No Change* : 0.2173, Fitur ini memberikan kontribusi 21.73% dan juga menjadi salah satu indikator utama. Artinya, informasi bahwa pasien sebelumnya tidak mengalami perubahan sangat membantu model dalam proses klasifikasi.
- j. Fitur dengan nilai 0.0000, seperti *Treatment Start Year*, *Outcome_Outcome*, dan beberapa placeholder kategorikal, tidak memberikan kontribusi apapun terhadap prediksi dan bisa diabaikan dalam analisis lebih lanjut.

BAB V

PENUTUP

5.1 Kesimpulan

Prediksi terhadap hasil pengobatan kesehatan mental merupakan sebuah proses analisis yang bertujuan untuk mengidentifikasi kecenderungan kondisi pasien setelah menjalani terapi, apakah menunjukkan kondisi membaik (*Improved*), memburuk (*Deteriorated*) atau tidak ada perubahan (*No Change*). Data yang digunakan dalam penelitian ini berasal dari dari pasien dengan berbagai atribut yang mencakup aspek klinis, emosional, dan perilaku. Data tersebut telah melalui tahapan preprocessing seperti pembersihan data, konversi tipe, pengkodean kategorikal (*one-hot encoding*), serta normalisasi. Total data yang dianalisis berjumlah 501 dan dibagi menjadi data pelatihan dan data pengujian dengan rasio 80:20.

Algoritma yang digunakan dalam penelitian ini adalah algoritma *Random Forest* untuk mengklasifikasi hasil pengobatan berdasarkan sejumlah fitur seperti usia, tingkat stres, tingkat kepatuhan terhadap pengobatan, skor mood, dan atribut lainnya. Hasil yang diperoleh menunjukkan bahwa algoritma *Random Forest* mampu menghasilkan klasifikasi dengan performa sangat tinggi, ditandai dengan akurasi mencapai 99,01%. Dengan nilai *precision*, *recall* dan *F1-score* tinggi untuk kelas *Improved* dan *Deteriorated* rata-rata 0,99. Fitur yang paling berpengaruh adalah riwayat pengobatan, usia, dan kepatuhan, yang menunjukkan pentingnya faktor historis dan perilaku dalam keberhasilan terapi.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, disarankan untuk penelitian selanjutnya melakukan perbandingan dengan algoritma lain atau menerapkan Teknik optimasi, sehingga dapat menghasilkan kinerja model yang lebih optimal.



DAFTAR PUSTAKA

- [1] F. Firmansyah and A. Yulianto, "Pemodelan Pembelajaran Mesin untuk Prediksi Kesehatan Mental di Tempat Kerja," *J. Minfo Polgan*, vol. 13, no. 1, pp. 397–407, 2024, doi: 10.33395/jmp.v13i1.13674.
- [2] Suryanto Aloysius and Nada Salvia, "Analisis Kesehatan Mental Mahasiswa Perguruan Tinggi Pada Awal Terjangkitnya Covid-19 di Indonesia," *J. Citizsh. Virtues*, vol. 1, no. 2, pp. 83–97, 2021.
- [3] L. Liesay, J. Mainase, and S. Yakobus, "Gambaran Gejala Gangguan Kesehatan Mental Berdasarkan Dass-42 (Depression Anxiety Stress Scales-42) Pada Masyarakat Usia Produktif Desa Hutumuri," *Molucca Medica*, vol. 16, no. 1, pp. 51–60, 2023, doi: 10.30598/molmed.2023.v16.i1.51.
- [4] C. T. Setiawan, S. G. Sijabat, Ervan, and Habibi, "Menjembatani Kesenjangan dalam Perawatan Kesehatan Mental: Pendekatan Baru untuk Diagnosis, Pengobatan, dan Pengurangan Stigma," *J. Multidisiplin West Sci.*, vol. 2, no. 08, pp. 660–667, 2023, doi: 10.58812/jmws.v2i08.579.
- [5] N. A. S. S. M. Utami Nur Hafsa Putri, *MODUL KESEHATAN MENTAL*. CV. AZKA PUSTAKA, 2022. [Online]. Available: https://books.google.co.id/books?id=yL_MEAAAQBAJ
- [6] 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.
- [7] A. Priyono, M. Shodiq, D. P. Alvinsyah, and S. A. Hidayah, "Metode Random Forest Untuk Memudahkan Klasifikasi Diagnosis Penyakit Mental," *J. Inform. Medis*, vol. 2, no. 1, pp. 1–4, 2024, doi: 10.52060/im.v2i1.2119.
- [8] S. S. A. Larasati, E. N. K. Dewi, B. H. Farhansyah, F. A. Bachtiar, and F. Pradana, "Penerapan Decision Tree dan Random Forest dalam Deteksi Tingkat Stres Manusia Berdasarkan Kondisi Tidur," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 10, no. 7, pp. 1503–1510, 2023, doi: 10.25126/jtiik.1077993.
- [9] S. P. M. M. Dr. Muhammad Ramdhan, *Metode Penelitian*. Cipta Media Nusantara. [Online]. Available: https://books.google.co.id/books?id=Ntw_EAAAQBAJ
- [10] S. K. M. K. Listra Frigia Missianes Horhoruw, S. S. T. M. T. Aditya Pratama, S. T. M. Yessy Asri, and S. K. M. K. Noviyanti. P, *EKSPLORASI MACHINE LEARNING DENGAN SCIKIT-LEARN Strategi Belajar Machine Learning*. Uwais Inspirasi Indonesia, 2024. [Online]. Available: <https://books.google.co.id/books?id=8GsWEQAAQBAJ>
- [11] G. Zahra, N. Fadhilah, R. A. Saputra, and A. H. Wibowo, "Deteksi Tingkat

- Gangguan Kecemasan Menggunakan Metode Random Forest,” *J. Fak. Tek. UMT*, vol. 13, no. 1, pp. 38–47, 2024, [Online]. Available: <http://jurnal.umt.ac.id/index.php/jt/index>
- [12] 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. Teknol.*, vol. 6, no. 2, p. 149, 2021, doi: 10.30998/string.v6i2.10361.
- [13] M. S. Joko Pramono. S. Sos., *Implementasi dan Evaluasi Kebijakan Publik*. Unisri Press, 2020. [Online]. Available: <https://books.google.co.id/books?id=yrubEAAAQBAJ>
- [14] D. Jollyta, A. Hajjah, E. Haerani, and M. Siddik, *Algoritma Klasifikasi untuk Pemula Solusi Python dan RapidMiner*. Deepublish, 2023. [Online]. Available: <https://books.google.co.id/books?id=y84TEQAAQBAJ>
- [15] M. A. Prof. Dr. H. M. Bahri Ghazali, *KESEHATAN MENTAL: Membangun Hidup Lebih Bermakna*. Samudra Biru, 2024. [Online]. Available: <https://books.google.co.id/books?id=eIsJEQAAQBAJ>
- [16] *Konsep Kesehatan Mental Zakiah Daradjat: Relevansinya dengan Kecerdasan Emosional dan Spiritual*. Penerbit NEM, 2022. [Online]. Available: <https://books.google.co.id/books?id=F9qgEAAAQBAJ>
- [17] S. S. M. K. Muhammad Arhami and S. T. M. T. Muhammad Nasir, *Data Mining - Algoritma dan Implementasi*. Andi Offset, 2020. [Online]. Available: <https://books.google.co.id/books?id=AtcCEAAAQBAJ>
- [18] S. K. M. K. D. A. N. A. P. S. K. M. K. Amril Mutoi Siregar, *DATA MINING: Pengolahan Data Menjadi Informasi dengan RapidMiner*. CV Kekata Group. [Online]. Available: <https://books.google.co.id/books?id=rTImDwAAQBAJ>
- [19] S. Junaidi *et al.*, *Buku Ajar Machine Learning*. PT. Sonpedia Publishing Indonesia, 2024. [Online]. Available: <https://books.google.co.id/books?id=ACT2EAAAQBAJ>
- [20] J. A. Nursiyono, *MACHINE LEARNING dengan R Teori & Praktikum*. Media Nusa Creative (MNC Publishing), 2023. [Online]. Available: <https://books.google.co.id/books?id=4T-xEAAAQBAJ>
- [21] R. F. Putra *et al.*, *DATA MINING : Algoritma dan Penerapannya*. PT. Sonpedia Publishing Indonesia, 2023. [Online]. Available: <https://books.google.co.id/books?id=zLHGEEAAAQBAJ>
- [22] R. N. Alifah *et al.*, “Perbandingan Metode Tree Based Classification untuk Masalah Klasifikasi Data Body Mass Index,” *Indones. J. Math. Nat. Sci.*, vol. 47, no. 1, p. 2024, 2024, [Online]. Available: <https://journal.unnes.ac.id/journals/JM/index>

- [23] A. T. Wahyudi, *Pemrograman Komputer menggunakan Flowgorithm dan APP inventor*. USB Press, 2021. [Online]. Available: <https://books.google.co.id/books?id=lqFKEAAAQBAJ>
- [24] A. D. P. Rusman and U. Suwardoyo, *Penerapan Sistem Informasi Berbasis IT Pengolahan Data Rekam Medis untuk Peningkatan Pelayanan di Rumah Sakit*. Penerbit NEM, 2022. [Online]. Available: <https://books.google.co.id/books?id=8XCAEAAAQBAJ>
- [25] *Dasar-Dasar Data Science dan Aplikasinya dengan Python*. wawasan Ilmu. [Online]. Available: <https://books.google.co.id/books?id=vbYOEQAAQBAJ>
- [26] E. Supriyadi, N. hikmah, M. Sofiana, and P. Prasetiyawati, *Panduan Praktis Belajar Microsoft Excel 2021*. in Edisi Pertama. TOHAR MEDIA, 2022. [Online]. Available: <https://books.google.co.id/books?id=ukWEEAAAQBAJ>
- [27] F. S. M. Kom, *Buku Ajar Dasar Pemrograman (Python)*. Nusa Putra Press, 2021. [Online]. Available: <https://books.google.co.id/books?id=zA08EAAAQBAJ>
- [28] S. Maesaroh *et al.*, *Bahasa Pemrograman Python*. Sada Kurnia Pustaka, 2024. [Online]. Available: <https://books.google.co.id/books?id=bOIKEQAAQBAJ>
- [29] B. Nurhakim, A. R. Dikananda, R. Hamonangan, R. A. Prasetyo, and A. Sofatunisa, *Data Cerdas di Dunia Pendidikan: Penerapan Machine Learning untuk Kepuasan Mahasiswa*. MEGA PRESS NUSANTARA, 2024. [Online]. Available: <https://books.google.co.id/books?id=Rb1HEQAAQBAJ>

CURRICULUM VITAE

Nama : Siti Romlah
NPM : 2021501061
Tempat, Tanggal Lahir : Situbondo, 21 September 2002
Program Studi : Sistem Informasi
Nama Orang Tua
Ayah : Sarim
Ibu : Sumiyati
Jenjang Pendidikan
SD/MI : SDIT Nurul Anshor
SMP/MTS : MTS At- Tamimi
SMK/MA : SMK Ibrahimi 1 Sukorejo
Latar Organisasi : -
Alamat Rumah : Bukit-Sema, Ardirejo, Kec. Panji, Kab. Situbondo
No. Telepon : 085338270349
E-mail : Sitiromla744@gmail.com

LAMPIRAN

Lampiran A : Data Mentah

| Patient ID | Age | Gender | Diagnosis | Symptom Severity (1-10) | Mood Score (1-10) | Sleep Quality (1-10) | Physical Activity (hrs/week) | Medication | Therapy Type | Treatment Start Date | Treatment Duration (weeks) | Stress Level (1-10) | Treatment Progress (1-10) | AI-Detected Emotional State | Adherence to Treatment (%) | Outcome |
|------------|-----|--------|---------------------------|-------------------------|-------------------|----------------------|------------------------------|------------------|--------------------------------|----------------------|----------------------------|---------------------|---------------------------|-----------------------------|----------------------------|--------------|
| 1 | 43 | Female | Major Depressive Disorder | 10 | 5 | 8 | 5 | Mood Stabilizers | Interpersonal Therapy | 25/01/2024 | 11 | 9 | 7 | Anxious | 66 | Deteriorated |
| 2 | 40 | Female | Major Depressive Disorder | 9 | 5 | 4 | 7 | Antipsychotics | Interpersonal Therapy | 27/02/2024 | 11 | 7 | 7 | Neutral | 78 | No Change |
| 3 | 55 | Female | Major Depressive Disorder | 6 | 3 | 4 | 3 | SSRIs | Mindfulness-Based Therapy | 20/03/2024 | 14 | 7 | 5 | Happy | 62 | Deteriorated |
| 4 | 34 | Female | Major Depressive Disorder | 6 | 3 | 6 | 5 | SSRIs | Mindfulness-Based Therapy | 29/03/2024 | 8 | 8 | 10 | Excited | 72 | Deteriorated |
| 5 | 52 | Male | Panic Disorder | 7 | 6 | 6 | 8 | Anxiolytics | Interpersonal Therapy | 18/03/2024 | 12 | 5 | 6 | Excited | 63 | Deteriorated |
| 6 | 28 | Male | Panic Disorder | 8 | 7 | 6 | 4 | SSRIs | Cognitive Behavioral Therapy | 11/01/2024 | 13 | 9 | 7 | Stressed | 82 | No Change |
| 7 | 59 | Male | Generalized Anxiety | 6 | 6 | 5 | 3 | Mood Stabilizers | Dialectical Behavioral Therapy | 21/02/2024 | 13 | 8 | 5 | Anxious | 79 | Deteriorated |
| 8 | 32 | Female | Major Depressive Disorder | 5 | 3 | 5 | 2 | Antidepressants | Cognitive Behavioral Therapy | 03/02/2024 | 12 | 5 | 6 | Stressed | 74 | No Change |
| 9 | 21 | Female | Major Depressive Disorder | 6 | 7 | 7 | 10 | Antidepressants | Mindfulness-Based Therapy | 16/01/2024 | 15 | 5 | 6 | Neutral | 67 | No Change |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 491 | 33 | Female | Bipolar Disorder | 10 | 3 | 7 | 1 | Mood Stabilizers | Interpersonal Therapy | 02/03/2024 | 16 | 10 | 8 | Happy | 83 | Deteriorated |
| 492 | 42 | Male | Panic Disorder | 10 | 3 | 7 | 10 | Mood Stabilizers | Cognitive Behavioral Therapy | 14/02/2024 | 13 | 8 | 6 | Anxious | 85 | Improved |
| 493 | 18 | Male | Bipolar Disorder | 10 | 5 | 8 | 7 | Antipsychotics | Mindfulness-Based Therapy | 28/01/2024 | 11 | 8 | 9 | Neutral | 71 | Improved |
| 494 | 23 | Male | Panic Disorder | 9 | 4 | 8 | 2 | Anxiolytics | Interpersonal Therapy | 04/01/2024 | 9 | 8 | 6 | Happy | 60 | Improved |
| 495 | 19 | Male | Major Depressive Disorder | 8 | 5 | 5 | 2 | SSRIs | Dialectical Behavioral Therapy | 05/02/2024 | 14 | 7 | 10 | Stressed | 70 | Improved |
| 496 | 24 | Male | Generalized Anxiety | 10 | 4 | 8 | 6 | Mood Stabilizers | Dialectical Behavioral Therapy | 09/04/2024 | 8 | 9 | 10 | Depressed | 73 | Improved |
| 497 | 22 | Male | Panic Disorder | 5 | 6 | 6 | 7 | Benzodiazepines | Mindfulness-Based Therapy | 05/02/2024 | 13 | 6 | 8 | Happy | 86 | Deteriorated |
| 498 | 23 | Male | Major Depressive Disorder | 7 | 3 | 4 | 2 | Antidepressants | Cognitive Behavioral Therapy | 24/03/2024 | 10 | 5 | 5 | Neutral | 87 | Deteriorated |
| 499 | 48 | Male | Bipolar Disorder | 9 | 4 | 6 | 9 | Antidepressants | Mindfulness-Based Therapy | 22/03/2024 | 10 | 6 | 7 | Anxious | 73 | Improved |
| 500 | 58 | Male | Bipolar Disorder | 5 | 7 | 6 | 2 | SSRIs | Cognitive Behavioral Therapy | 24/03/2024 | 10 | 6 | 5 | Happy | 70 | Deteriorated |

Lampiran B : Data Preprocessing

| Index | Age | Sex | Marital | Education | Income | Religion | Occupation | Smoking | Alcohol | Depression | Medication | Therapy | Adherence | Outcome |
|-------|-----|--------|---------|-------------|--------|----------|------------|------------|---------|------------------|-----------------|--------------------|-----------|----------|
| 1 | 40 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 2 | 40 | Female | Single | High School | Low | Islam | Unemployed | Non-smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 3 | 35 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 4 | 34 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 5 | 52 | Female | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 6 | 30 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 7 | 59 | Female | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 8 | 32 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 9 | 21 | Female | Single | High School | Low | Islam | Unemployed | Non-smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 10 | 33 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 11 | 42 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 12 | 38 | Female | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 13 | 49 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 14 | 37 | Female | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 15 | 45 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 16 | 31 | Female | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 17 | 41 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 18 | 36 | Female | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 19 | 43 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 20 | 39 | Female | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 21 | 47 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 22 | 34 | Female | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 23 | 46 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 24 | 32 | Female | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 25 | 44 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 26 | 38 | Female | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 27 | 48 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 28 | 35 | Female | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 29 | 42 | Male | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |
| 30 | 37 | Female | Married | High School | Low | Islam | Unemployed | Smoker | Alcohol | Major Depression | Antidepressants | Behavioral Therapy | 75% | Improved |

Lampiran D : Hasil Cek Plagiasi



PONDOK PESANTREN SALAFIYAH SYAFI'YAH SUKOREJO
UNIVERSITAS IBRAHIMI
PERPUSTAKAAN IBRAHIMI
N P P . 3 5 1 2 1 4 2 F 2 0 0 6 5 6 7
Jl. KHR. Syamsul Arifin No. 1-2 PO. Box. 2 Kode Pos. 68374 Phone (0338) 452666 Fax. (0338) 453068
SUMBEREJO BANYUPUTIH SITUBONDO JAWA TIMUR

SURAT KETERANGAN
HASIL PEMERIKSAAN PLAGIASI

Yang bertanda tangan di bawah ini

Nama : Muhammad Ali Ridla, M.Kom.
Jabatan : Kepala Perpustakaan

Menyatakan dengan sebenarnya bahwa:

NIM : 2021502061
Nama : SITI ROMLAH
Fakultas : Sains dan Teknologi
Prodi : Sistem Informasi
Kecamatan : PANJI
Kabupaten : SITUBONDO
Provinsi :
Judul Skripsi : Implementasi Algoritma Random Forest Untuk
Memprediksi Hasil Pengobatan Kesehatan Mental

Dengan dosen Pembimbing :

1. Firman Santoso, M.Kom.
2. Nur Azise, M.Kom.

Telah dilakukan cek plagiasi di Perpustakaan Universitas Ibrahimi dengan persentase plagiasi terakhir sebesar 26% .

Demikian Surat Keterangan ini dibuat untuk dipergunakan sebagaimana mestinya.

Sukorejo, 6 Agustus 2025
Kepala Perpustakaan,



Muhammad Ali Ridla, M.Kom.



UU ITE No.11 Tahun 2008 Pasal 5 Ayat 1
"Informasi Elektronik dan/atau Dokumen Elektronik
dan/atau hasil cetaknya merupakan alat bukti yang sah."

Lampiran E : Lembar Pernyataan Kesiediaan Publikasi**LEMBAR PERNYATAAN
KESEDIAAN PUBLIKASI KARYA ILMIAH**

Saya yang bertanda tangan di bawah ini:

Nama : **SITI ROMLAH**
NIM/NPM : 2021502061
Program Studi : Sistem Informasi
Fakultas : Sains dan Teknologi
Jenis Karya Ilmiah : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan Hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty-Free Right*) kepada Perpustakaan Universitas Ibrahimi atas karya ilmiah saya berupa Skripsi yang berjudul:

**“IMPLEMENTASI ALGORITMA RANDOM FOREST UNTUK
MENGKLASIFIKASI HASIL PENGOBATAN KESEHATAN MENTAL”**

Dengan Hak Bebas Royalti Non-eksklusif ini Pusat Perpustakaan Universitas Ibrahimi berhak menyimpan, alih media/format, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat untuk dapat dipergunakan sebagaimana mestinya.

Situbondo, 19 Agustus 2025
Yang Menyatakan



METERAI
TEMPEL
BB 199AMX399085798

SITI ROMLAH