

# BAB I

## PENDAHULUAN

### 1.1. LATAR BELAKANG MASALAH

Gangguan kesehatan mental merupakan salah satu masalah kesehatan global yang semakin mendapatkan perhatian serius. Gangguan kesehatan mental seperti depresi, kecemasan, bipolar, dan *schizophrenia* dapat memiliki dampak yang signifikan pada kualitas hidup individu serta masyarakat pada umumnya. Menurut data *World Health Organization* (WHO) regional Asia Pasifik, jumlah penderita gangguan mental di Indonesia sebanyak 9.162.886 kasus.

Diagnostik gangguan kesehatan mental seringkali merupakan tantangan yang kompleks bagi profesional medis karena beragamnya gejala dan faktor yang terlibat. Dalam konteks ini, penggunaan teknologi dan analisis data dapat sangat membantu dalam diagnosis gangguan kesehatan mental. *Data mining* adalah ekstraksi informasi atau pola yang penting atau menarik dari data yang ada di database yang besar [1]. Dalam data mining terdapat beberapa metode, salah satunya algoritma C4.5.

Algoritma C4.5 merupakan algoritma yang digunakan untuk memprediksi sebuah keputusan dengan menerapkan serangkaian aturan keputusan [2]. Algoritma C4.5 tidak hanya ditentukan dengan perhitungan manual, namun bisa juga dengan menggunakan aplikasi Rapid Miner dan Google Colaboratory. Rapid Miner adalah *platform* perangkat lunak ilmu data yang menyediakan lingkungan terintegrasi untuk persiapan data, pembelajaran mesin, pembelajaran dalam,

penambahan teks, dan analisis prediktif [3]. Sedangkan Google Colab adalah sebuah IDE untuk pemrograman *Python* dimana pemrosesan akan dilakukan oleh server Google yang memiliki perangkat keras dengan performa yang tinggi [4].

Beberapa penelitian telah dilakukan untuk memprediksi gangguan kesehatan mental menggunakan metode Algoritma C4.5 berjudul Analisis Data Hasil Diagnosa Untuk Klasifikasi Gangguan Kepribadian Menggunakan Algoritma C4.5. oleh Siska Febriani dan Heni Sulistiani dengan hasil nilai akurasi data sebesar 72.67% [5]. Penelitian kedua berjudul Pemodelan Prediksi Kesehatan Mental Mahasiswa Di Lingkungan Multikultural Menggunakan Algoritma Decision Tree oleh Ahmad Haidir dengan hasil nilai akurasi sebesar 84.96% [6]. Penelitian sebelumnya telah menemukan akurasi yang sangat baik, tetapi masih ada ruang untuk perbaikan. Tujuan dari penelitian ini adalah untuk mengevaluasi penerapan Algoritma C4.5 dalam memprediksi gangguan kesehatan mental menggunakan aplikasi Rapid Miner dan Google Colaboratory.

## **1.2. RUMUSAN MASALAH**

Berdasarkan latar belakang, maka masalah yang dapat dirumuskan adalah:

1. Bagaimana mengimplementasi Algoritma C4.5 untuk prediksi gangguan kesehatan mental menggunakan Rapid Miner dan Google Colaboratory?

2. Bagaimana mengevaluasi penerapan Algoritma C4.5 untuk prediksi gangguan kesehatan mental menggunakan Rapid Miner dan Google Colaboratory?

### 1.3. BATASAN MASALAH

Untuk menghindari pembahasan yang meluas pada penelitian ini, maka penulis memberikan batasan masalah yaitu:

1. Prediksi gangguan kesehatan mental pada penelitian ini hanya menggunakan Algoritma C4.5.
2. Penentuan Algoritma C4.5 pada penelitian ini hanya menggunakan aplikasi Rapid Miner dan Google Colaboratory dengan bahasa pemrograman Python.
3. Dataset yang akan digunakan pada penelitian ini adalah data yang diambil dari Kaggle yaitu "*Unemployment and Mental Illness Survey*" oleh Michael Corley, MBA, LSSBB, CPM dengan jumlah data sebanyak 335 data (sumber: <https://www.kaggle.com/datasets/michaelcorley/unemployment-and-mental-illness-survey/data>).
4. Variabel pada dataset yang akan digunakan untuk melakukan klasifikasi antar lain *I have been hospitalized before for my mental illness, How many days were you hospitalized for your mental illness, How many times were you hospitalized for your mental illness, Lack of concentration, Anxiety, Depression, Obsessive thinking, Mood swings,*

5. *Class/target* pada dataset ini adalah variabel *I identify as having a mental illness* dengan tipe data *binominal (Yes, No)*.
6. Hasil penelitian ini akan menampilkan evaluasi dari *cross validation*, dan *confussion matrix*, serta variabel yang paling berpengaruh dalam memprediksi gangguan kesehatan mental.

#### **1.4. TUJUAN DAN MANFAAT PENELITIAN**

##### **1.4.1. Tujuan Penelitian**

Adapun tujuan dari penelitian ini adalah:

1. Mengimplementasi Algoritma C4.5 dalam memprediksi gangguan kesehatan mental menggunakan Rapid Miner dan Google Colaboratory.
2. Mengevaluasi penerapan Algoritma C4.5 dalam memprediksi gangguan kesehatan mental menggunakan Rapid Miner dan Google Colaboratory.

##### **1.4.2. Manfaat Penelitian**

Manfaat dari Implementasi Algoritma C4.5 Untuk Prediksi Penyakit Mental Menggunakan Rapid Miner dan Google Colaboratory antara lain:

1. Dengan menggunakan algoritma C4.5, penelitian ini dapat membantu meningkatkan akurasi prediksi gangguan kesehatan mental. Hal ini

dapat membantu dalam diagnosis dan pengobatan gangguan kesehatan mental yang lebih efektif.

2. Rapid Miner dan Google Colaboratory dapat membantu meningkatkan efisiensi dalam pengolahan data. Hal ini dapat membantu dalam pengambilan keputusan yang lebih cepat dan akurat.
3. Dengan menganalisis faktor-faktor yang berkontribusi terhadap gangguan kesehatan mental, penelitian ini dapat memberikan wawasan yang lebih baik tentang penyebab dan karakteristik gangguan kesehatan mental.
4. Dengan memahami faktor-faktor yang memengaruhi gangguan kesehatan mental, penelitian ini dapat membantu dalam merancang alat atau program pemantauan yang lebih efektif untuk masyarakat yang memerlukan perhatian kesehatan mental.

## **1.5. SISTEMATIKA PENULISAN**

Sistematika penulisan merupakan kerangka dan pedoman penulisan tugas akhir yang berisi uraian singkat mengenai pembahasan yang ada pada penelitian ini secara menyeluruh. Adapun sistematika penulisannya adalah sebagai berikut:

### **BAB I            PENDAHULUAN**

Bab ini terdiri dari latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan tugas akhir.

## BAB II LANDASAN TEORI

Bab ini berisi teori-teori dasar yang mendukung penelitian. Dalam penelitian ini, teori-teori yang dimuat antara lain implementasi, algoritma C4.5, klasifikasi, gangguan kesehatan mental, Rapid Miner, Google Colaboratory, evaluasi, *confussion matrix* dan *cross validation*. Landasan teori juga memuat tinjauan penelitian sejenis dalam 5 tahun terakhir.

## BAB III METODOLOGI PENELITIAN

Bab ini berisi metode atau pendekatan yang akan digunakan dalam menjawab permasalahan penelitian untuk mencapai tujuan penelitian. Dalam penelitian ini, metodologi penelitian berisi alat dan bahan penelitian, metode penelitian, kerangka kerja penelitian, serta jadwal penelitian.

## BAB IV ANALISIS

Bab ini berisi analisis hasil penelitian yang dilakukan dengan menggunakan *tool* Rapid Miner dan Google Colaboratory. Beberapa hal yang disajikan dalam bab ini adalah profil data, analisis persiapan data, kecocokan algoritma C4.5 dengan karakteristik data, alasan pemilihan

*tool* Rapid Miner, hasil analisis data, serta kesimpulan dari hasil interpretasi data.

## BAB V KESIMPULAN

Berisi kesimpulan dan saran dari hasil penelitian secara keseluruhan. Kesimpulan yang dipaparkan harus konsisten dengan tujuan penelitian serta harus mencerminkan terpecahkan atau tidak masalah yang dibahas pada bab I.