

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada saat ini dunia sedang waspada dengan adanya penyebaran sebuah virus yang populer yaitu virus corona. Virus corona adalah bagian dari keluarga virus yang disebabkan penyakit mulai dari flu. Penyakit yang disebabkan virus corona ini dikenal dengan sebutan Covid-19 yaitu penemuan jenis baru tahun 2019 dan belum pernah diidentifikasi sebelumnya menyerang manusia. Pertama kali virus corona muncul dan menyerang manusia di provinsi Wuhan, China. Gejala virus corona yaitu batuk, sesak nafas, lelah, demam, dan tidak nafsu makan. Virus corona berkembang dengan sangat cepat sehingga dapat mengakibatkan infeksi yang lebih parah bahkan gagal organ. Kondisi ini umumnya terjadi pada pasien yang sebelumnya memiliki masalah kesehatan. Covid-19 telah menyebar ke 196 Negara, kasus yang terkonfirmasi positif ada sebanyak 414,179 dan kasus yang terkonfirmasi meninggal dunia sebanyak 18,440 [1].

Organisasi Kesehatan Dunia (WHO) menetapkan wabah virus corona sebagai pandemi pada 11 Maret 2020. Wabah virus corona ini sangat mengguncang masyarakat dunia hampir 200 Negara di dunia terjangkit oleh virus ini. Bentuk upaya pencegahan penyebaran virus Covid-19 dilakukan oleh pemerintah di negara-negara di dunia dengan cara memutus rantai penyebaran virus Covid-19 ini, yang dikenal dengan *lockdown dan social distancing*. Dengan adanya cara ini setiap masyarakat tidak akan menjadi penular maupun tertular karena tidak melakukan kontak dengan siapapun sehingga laju penyebaran dapat menurun [2]. Tetapi tidak menutup kemungkinan bahwa penularan ini dapat terjadi baik untuk diri sendiri maupun untuk orang-orang di sekitar kita, termasuk keluarga.

Pada penelitian [3] menjelaskan tentang algoritma C4.5 dalam mendiagnosis dan mengklasifikasikan penyakit, seperti penelitian tentang analisis rekam medis untuk menentukan pola kelompok penyakit menggunakan algoritma C4.5. Hasil perhitungan algoritma C4.5 mampu menganalisis trend penyakit yang dialami masyarakat.

Kemudian ada penelitian tentang algoritma C4.5 untuk diagnosis penyakit pneumonia. Hasil dari penelitian tersebut menunjukkan bahwa algoritma C4.5 berhasil memodelkan pohon keputusan dengan 10 aturan pneumonia [4].

Pada Penelitian [5] menjelaskan tentang pengklasifikasian dataset surveilans Covid-19 menggunakan *decision tree*. Dataset surveilans Covid-19 ini didapatkan dari portal data publik yaitu UCI machine learning repository. Penelitian ini menghasilkan akurasi yang lebih baik dengan tingkat akurasi yang diperoleh dengan menggunakan algoritma *decision tree* adalah 65%. Pada penelitian ini tingkat akurasi masih tergolong rendah. Rendahnya tingkat akurasi ini disebabkan dataset yang digunakan hanya memiliki 7 atribut dan 14 record.

Pada Penelitian [6] Penelitian dilakukan untuk memprediksi hasil belajar siswa pada masa pandemi Covid-19 sehingga hasil penelitian bisa digunakan untuk menjadi acuan dalam pengambilan kebijakan di sekolah. Data nilai siswa yang diolah diambil dari 1 (satu) mata pelajaran sebagai sampel data penelitian. Nilai hasil belajar siswa dianalisa menggunakan Metode C4.5 untuk mendapatkan pengetahuan baru dari data nilai hasil belajar siswa yang dilaksanakan pada masa pandemi Covid-19. Data yang dianalisa terdiri dari atribut absensi, tugas, ulangan harian dan nilai ujian yang mempengaruhi kriteria keputusan hasil belajar siswa pada pelajaran yang dilaksanakan secara daring. Kriteria keputusan hasil belajar terdiri dari “Memuaskan” dan “Kurang Memuaskan”.

Dari penjelasan diatas terdapat beberapa permasalahan yang harus diatasi dalam pencegahan penyebaran atau penularan virus Covid-19 ini. Kurangnya pengetahuan dan kesadaran masyarakat luas dalam melaksanakan pencegahan dan pengendalian Covid-19 adalah salah satu bentuk faktor meningkatnya penyebaran virus Covid-19. Kemudian ada kesulitan dalam melakukan pengklasifikasian penderita Covid 19, identifikasi dini atau diagnosa lebih awal, pencegahan dan pengendalian infeksi, komunikasi resiko dan pemberdayaan masyarakat. Beberapa masalah tersebut disebabkan oleh kurangnya implementasi dan pengujian metode kecerdasan buatan untuk mendiagnosis virus Covid-19 yang dapat digunakan secara publik. Diagnosis ini dapat dioptimalkan, diimplementasikan, dan diuji

menggunakan salah satu metode kecerdasan buatan yaitu algoritma *decision tree* serta dengan bantuan *website*. Dikarenakan *dataset* virus Covid-19 yang digunakan banyak maka dibutuhkan *machine learning* (mesin pembelajaran) dalam menyusun pohon keputusan. Dengan memasukkan data kemudian kita akan meminta mesin (*machine learning*) untuk membuat pohonnya. Kemudian untuk memudahkan seseorang dalam mengidentifikasi tingkat keparahan virus Covid-19 nya digunakan sebuah *website* yang dapat diakses secara umum.

Hasil penelitian ini ditujukan untuk meningkatkan ketersediaan fasilitas yang dapat digunakan untuk mengidentifikasi virus covid-19 secara cepat dan efisien. Sehingga penanganan pada pasien terdampak akan lebih efisien, serta keberadaan sistem ini juga dapat menutup kekurangan metode pengujian RT-PCR. Selain itu, hasil penelitian ini memberikan peluang untuk dapat diimplementasikan di daerah yang memiliki jumlah tenaga kesehatan terbatas, sehingga dapat menjadi solusi dalam mengatasi jumlah pasien yang terpapar virus Covid-19 yang terus meningkat. Dengan adanya *website* ini juga dapat meningkatkan kualitas kerja dan keakuratan dalam penyajian informasi sehingga pengguna tidak memiliki batasan waktu untuk mengakses informasi yang disajikan tetapi harus memiliki jaringan internet untuk mengaksesnya.

Dari permasalahan yang telah diuraikan sebelumnya, penelitian ini bertujuan untuk membantu mengidentifikasi tingkat keparahan virus Covid-19 seseorang berdasarkan beberapa variabel. Adapun dataset Covid-19 dalam penelitian ini diambil dari *website* kaggle. Kaggle adalah situs dan platform untuk membuat model terbaik dalam menganalisa dan memprediksi suatu dataset. Data kaggle adalah data publik yang telah teruji di segala bidang. Untuk bidang kesehatan dan ada beberapa data, termasuk data kasus Covid-19 yang saat ini menjadi tren penelitian dalam beberapa tahun terakhir. Berdasarkan hal tersebut, maka penulis melakukan sebuah penelitian dengan judul **“Implementasi Algoritma Decision Tree Machine Learning dalam Mengidentifikasi Penderita Virus Covid-19 Menggunakan Dataset Publik”**.

1.2 Rumusan Masalah

Adapun berdasarkan latar belakang masalah yang telah diuraikan sebelumnya, maka penulis mengemukakan beberapa rumusan masalah sebagai berikut :

1. Bagaimana cara membuat sebuah model dan akurasi *decision tree machine learning* dalam mengambil keputusan dan mengklasifikasikan penderita virus Covid-19 ?
2. Bagaimana menerapkan model *decision tree machine learning* untuk mengidentifikasi tingkat keparahan penderita virus Covid-19 yang dapat diintegrasikan ke dalam *website* ?

1.3 Ruang Lingkup dan Batasan Masalah

Untuk membatasi ruang lingkup permasalahan yang akan dibahas, maka dalam penulisan tugas akhir ini penulis lebih menekankan pada :

1. Dataset yang digunakan adalah penderita Covid-19 dengan enam variabel utama yaitu negara, usia, gejala, mengalami gejala lain, tingkat keparahan dan kontak.
2. Mengidentifikasi seseorang mengidap penyakit Covid-19 dengan tingkat keparahan berdasarkan beberapa gejala standar dan gejala lain yang telah ditentukan yaitu demam, kelelahan, kesulitan bernafas, batuk kering, dan sakit tenggorokan. Kemudian untuk gejala lain yaitu nyeri, hidung tersumbat, hidung beringsus, diare dan tidak bergejala.
3. Metode pengambilan keputusan menggunakan metode *decision tree machine learning*.
4. Bahasa pemrograman yang digunakan adalah bahasa pemrograman *Python*.
5. Tingkat performansi dan keakurasian dari sistem yang dibangun menggunakan *machine learning* dengan bahasa pemrograman *Python*.
6. Implementasi hasil model *decision tree* ke *website* dengan *framework flask*

1.4 Tujuan Penelitian

Atas dasar perumusan masalah, maka tujuan dalam perancangan alat adalah sebagai berikut :

1. Untuk mengetahui cara membuat sebuah model dan akurasi *decision tree machine learning* dalam pengambilan keputusan dan mengklasifikasikan penderita Covid-19 sehingga dapat diimplementasikan kedalam sebuah *website*.
2. Dapat mengetahui tentang metode *decision tree machine learning* kemudian mengetahui prosedur integrasi data
3. Dapat memahami proses perancangan *website*.

1.5 Manfaat Penelitian

Adapun manfaat yang akan didapatkan dalam penelitian ini adalah sebagai berikut :

1. Dapat memberikan identifikasi dini atau diagnosa lebih awal dalam mengidentifikasi Covid-19 pada seseorang.
2. Dengan adanya pengklasifikasian penderita Covid-19 tersebut, dapat dijadikan pencegahan dan pengendalian awal supaya virusnya tidak menyebar.
3. Dapat membantu mengidentifikasi tingkat keparahan seseorang yang mengidap penyakit Covid-19
4. Memudahkan tenaga medis dalam persiapan perencanaan dalam mengatasi dan mencegah virus Covid-19.

1.6 Metodologi Penulisan

Adapun metode penulisan dalam penelitian yang digunakan adalah sebagai berikut :

1. Metode Konsultasi atau Wawancara
Yaitu metode yang dilaksanakan melalui tanya jawab secara langsung dengan dosen pembimbing dan pihak lain yang berhubungan dengan

tugas akhir penulis.

2. Metode Studi Pustaka

Yaitu metode pengumpulan data mengenai fungsi dan cara kerja metode *decision tree* serta komponen lainnya yang bersumber baik dari buku, internet, jurnal dan lain sebagainya. Metode ini dilakukan untuk membantu Penulis dalam pembuatan Tugas Akhir.

3. Metode Perancangan

Yaitu metode dengan tahapan perancangan sistem berdasarkan dari hasil studi pustaka, pemodelan dari sistem tersebut diterjemahkan ke program simulasi dengan software *Python*.

4. Metode Observasi

Yaitu metode pengamatan terhadap data penyakit virus Covid-19 dari situs yang terkenal di dunia *data science* dan *machine learning* yaitu Kaggle sebagai acuan pengambilan informasi.

5. Metode Cyber

Yaitu metode yang dilakukan dengan cara mencari informasi dan data yang ada kaitannya dengan masalah yang dibahas dari internet sebagai bahan referensi Tugas Akhir.

1.7 Sistematika Penulisan

Dalam penulisan tugas akhir ini, sistematika penulisan terdiri dari beberapa bab dengan perincian sebagai berikut :

BAB I PENDAHULUAN

Pada bab ini menjelaskan tentang latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat, metodologi penulisan, serta sistematika penulisan.

BAB II TINJAUAN UMUM

Pada bab ini menjelaskan tentang teori-teori yang mendukung dan mendasari komponen yang akan digunakan, serta perbandingan dari penelitian sebelumnya.

BAB III METODOLOGI PENELITIAN

Pada bab ini menjelaskan metode dan proses perancangan sistem yang akan dibuat seperti blok-blok diagram, langkah kerja dan prinsip kerja sistem.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini menjelaskan hasil yang akan dicapai dengan menggunakan metodologi yang telah ditentukan sebelumnya. Bab ini juga merencanakan waktu yang dibutuhkan dalam perancangan sistem.

BAB V KESIMPULAN DAN SARAN

Bab ini merupakan bagian akhir dari penulisan tugas akhir yang berisi tentang kesimpulan serta saran.