

BAB I

PENDAHULUAN

1.1. Latar Belakang

Coronavirus adalah suatu kelompok virus yang dapat menyebabkan penyakit pada hewan atau manusia. Beberapa jenis coronavirus diketahui menyebabkan infeksi saluran nafas pada manusia mulai dari batuk pilek hingga yang lebih serius seperti *Middle East Respiratory Syndrome (MERS)* dan *Severe Acute Respiratory Syndrome (SARS)*. Coronavirus jenis baru yang ditemukan menyebabkan penyakit COVID-19. Virus baru dan penyakit yang disebabkan ini tidak dikenal sebelum mulainya wabah di Wuhan, Tiongkok, bulan Desember 2019. COVID-19 ini sekarang menjadi sebuah pandemi yang terjadi di banyak negara di seluruh dunia (WHO, 2021).

Saat ini ada sebanyak 65 negara terinfeksi virus corona. Menurut data WHO per tanggal 2 Maret 2020 jumlah penderita 90.308 terinfeksi Covid-19. Angka kematian mencapai 3.087 atau 2.3% dengan kesembuhan 45.726 orang (WHO, 2020). Situasi COVID-19 di Indonesia kurang lebih 130.718 kasus positif dengan 85.798 kasus sembuh dan 5.903 kematian hingga Agustus 2020 yang telah menyebar di 34 provinsi di Indonesia. Penderita COVID-19 terbanyak berada di DKI Jakarta (26.624 kasus), Jawa Timur (25.917 kasus), Jawa Tengah (10.765 kasus) (Harahap, 2020).

Infeksi COVID-19 dapat menimbulkan gejala ringan, sedang atau berat. Gejala klinis utama yang muncul yaitu demam (suhu > 38C), batuk dan kesulitan bernapas. Selain itu dapat disertai dengan sesak memberat, fatigue, mialgia, gejala gastrointestinal seperti diare dan gejala saluran napas lain. Setengah dari pasien timbul sesak dalam satu minggu. Pada kasus berat perburukan secara cepat dan progresif, seperti ARDS, syok septic, asidosis metabolik yang sulit dikoreksi dan pendarahan atau disfungsi sistem koagulasi dalam beberapa hari (Yuliana, 2020). Infeksi COVID-19 menyebar dengan cepat sehingga terjadi peningkatan jumlah pasien yang terinfeksi di seluruh dunia. Kasus COVID-19 yang serius dapat berkembang menjadi pneumonia berat, ARDS dan kegagalan beberapa organ

yang bisa menyebabkan kematian, sedangkan kasus tidak parah menunjukkan gejala biasa infeksi sistem pernapasan (Harahap, 2020).

Berdasarkan latar belakang diatas, penulis melakukan penelitian berupa sistem yang dapat mengidentifikasi penyakit COVID-19 ini dengan menggunakan algoritma *C5.0* yang didalamnya terdapat tahapan *pruning* berdasarkan gejala umum yang dialami pasien. Algoritma *C5.0* adalah salah satu algoritma klasifikasi data mining yang khususnya diterapkan pada teknik *decision tree*. *C5.0* merupakan penyempurnaan algoritma sebelumnya yang dibentuk oleh *Ross Quinlan* pada tahun 1987, yaitu *ID3* dan *C4.5*. Perbedaan dan kelebihan dari algoritma *C5.0* ini sendiri yaitu, adanya proses *boosting* didalamnya. *Boosting* pada algoritma *C5.0* merupakan proses pengacakan data latih yang tujuannya adalah untuk meningkatkan akurasi dari pohon keputusan yang telah dibentuk dari perhitungan algoritma *C5.0*.

Sedangkan untuk *pruning* pada *C5.0* sendiri merupakan salah satu cara untuk merampingkan struktur dari *decision tree* sehingga generalisasi datanya menjadi lebih baik, dan memudahkan interpretasinya. Cara kerja *pre-pruning* adalah dengan menghitung terlebih dahulu nilai *information gain* untuk mengetahui nilai *parent* dan *child*. Setelah *parent* dan *child* diketahui kemudian dihitung nilai erornya, jika nilai eror *child* lebih kecil dari *parent* maka *parent* membentuk *subtree* lagi, tapi sebaliknya jika nilai eror *child* lebih besar dari *parent* maka *pruning* dilakukan dan pembentukan *subtree* berhenti (Aric, Intan, & Fetty, 2019). Keluaran yang dihasilkan pada dalam penelitian ini yaitu sistem yang dapat mengidentifikasi penyakit COVID-19 dengan menggunakan algoritma *C5.0* yang ditanamkan didalam sistemnya.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijabarkan, maka didapatkan rumusan masalah yang akan dibahas antara lain :

- a. Bagaimana penerapan algoritma *C5.0* yang didalamnya terdapat tahapan *pruning* untuk mengidentifikasi penyakit covid-19 pada manusia berdasarkan gejala umum pada pasien?
- b. Bagaimana hasil evaluasi performa dari algoritma *C5.0* pada data uji?

- c. Bagaimana merancang dan membangun sistem yang menggunakan algoritma C5.0 untuk mengidentifikasi adanya penyakit covid-19?

1.3. Tujuan

Tujuan dari penelitian ini adalah membuat sistem yang dapat mendiagnosis penyakit covid-19 dengan menggunakan algoritma C5.0 dengan menerapkan metode Pruning berdasarkan input gejala yang diderita oleh pengguna. Sehingga mampu memudahkan pengguna untuk dapat mendeteksi penyakit covid lebih dini hanya dari gejala yang dialami oleh pengguna.

1.4. Manfaat

Dengan tercapainya tujuan penelitian diatas, maka diharapkan akan bermanfaat bagi semua pihak, manfaat dari penelitian ini adalah :

- a. Memberikan informasi diagnosis penyakit covid-19 berdasarkan gejala - gejala umum yang dialaminya.
- b. Mengetahui bagaimana algoritma C5.0 yang terdapat *pruning* didalamnya dapat diterapkan untuk melakukan identifikasi penyakit covid-19.

1.5. Batasan Masalah

Adapun batasan masalah yang digunakan Penulis agar pembahasan dalam penelitian ini tidak menyimpang dari pembahasan adalah sebagai berikut :

- a. Sumber data yang digunakan pada penelitian ini adalah data yang diambil dari dataset *Kaggle*.
- b. Data yang diolah hanya data yang berhubungan dengan gejala-gejala umum penyakit covid-19.
- c. Data yang digunakan adalah data excel.
- d. Algoritma yang akan digunakan untuk mengidentifikasi penyakit covid-19 adalah algoritma C5.0 yang didalamnya terdapat tahapan *pruning*.
- e. Keluaran yang dihasilkan pada penelitian ini berupa hasil analisis evaluasi performa pengujian algoritma C5.0 yang ditanamkan pada sistem.