

## BAB I

### PENDAHULUAN

#### 1.1 Latar Belakang

Coronavirus Disease (COVID-19) disebabkan oleh virus Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2). Virus ini diketahui dapat menular dari manusia ke manusia dan telah menyebar di China lebih dari 190 negara di dunia termasuk Indonesia. Pada Maret 2020 sendiri WHO juga telah mengumumkan COVID 19 sebagai pandemi global di seluruh dunia. Terhitung hingga tanggal 1 November 2021 telah terkonfirmasi jumlah kasus positif COVID 19 di seluruh dunia sebanyak 246.594.191 dengan total 4.998.784 kematian (WHO, 2021).

Jumlah pasien terkonfirmasi COVID 19 terus bertambah setiap harinya karena penyebaran virus COVID 19 yang sangat cepat, hal ini dikonfirmasi WHO dapat terjadi apabila seseorang yang sehat melakukan kontak langsung dengan seorang pasien yang telah terinfeksi virus tersebut, dimana virus ini dapat menyebar melalui percikan droplet yang disebarkan pada saat batuk atau bersin dari individu yang telah terinfeksi virus (Rothan and Byrareddy, 2020).

Gejala klinis yang biasanya muncul pada pasien penderita COVID 19 antara lain adalah demam, batuk kering dan sesak napas. Gejala lain yang biasanya timbul selain yang menyerang pernapasan adalah gejala yang menyerang pencernaan, dengan beberapa penelitian ditemukan bahwa pasien penderita COVID 19 juga mengalami sakit abdominal, mengalami diare, mual dan muntah (Aditia, 2021).

Kasus COVID 19 pertama di Indonesia, dilaporkan pada Maret 2020 dengan temuan sebanyak 2 kasus. Namun, kasus ini terus mengalami peningkatan setiap harinya. Hingga November 2021, telah tercatat jumlah kasus terkonfirmasi COVID 19 di Indonesia sebanyak 4.244.761 dengan total kasus kematian sebanyak 143.423 (Satgas covid 19, 2021).

Dilihat dari data kasus COVID 19 tersebut, Indonesia menjadi salah satu dari sekian banyak negara yang jumlah kasus pasien terkonfirmasi COVID 19 terus meningkat. Selain vaksin yang belum merata, faktor deteksi COVID 19 yang belum

optimal juga menjadi penyebab kasus yang terus bertambah. Terlepas dari pencegahan penularan virus yang telah diterapkan seperti karantina mandiri dan *social distancing*, WHO merekomendasikan tes untuk mendeteksi adanya virus COVID 19 di tubuh seseorang. Standar tes yang saat ini untuk deteksi SARS-CoV-2 adalah SARS-CoV-2 *real time reverse transcription quantification polimerase chain reaction* (RT-PCR) menggunakan sampel bahan swab nasofaring atau orofaring, sputum atau cairan bilas bronkial (*bronkial lavage*) dan rapid test (Pusparini, 2020). Selain menggunakan RT-PCR, rapid test antibody dapat digunakan juga sebagai salah satu metode yang dapat digunakan untuk mendeteksi COVID 19.

Di Indonesia sendiri, rapid test antibody digunakan sebagai pemeriksaan skrining adanya antibodi terhadap orang tanpa gejala (OTG) atau kasus kontak dari pasien konfirmasi COVID-19. Namun, pendeteksian ini hanya dapat diinterpretasikan ketika virus aktif bereplikasi, oleh sebab itu, alat ini paling tepat digunakan saat fase akut infeksi. Akan tetapi alat ini memiliki kekurangan yaitu, seseorang yang telah mendapatkan hasil dari tes tersebut tetap harus melakukan swab tes atau RT-PCR jika diketahui hasilnya reaktif.

Menurut WHO, RT-PCR memiliki tingkat keakuratan yang lebih tinggi dibandingkan dengan rapid test. Namun, RT-PCR ini memerlukan waktu pengerjaan yang cukup lama jika dibandingkan dengan rapid test, peralatan dan biaya pemeriksaan juga lebih mahal. Diantara kedua tes tersebut terdapat kelemahan yang sama yaitu, menimbulkan risiko pemaparan virus terhadap orang lain, dikarenakan untuk melakukan tes tersebut, seseorang harus pergi ke rumah sakit. Oleh karena itu, diperlukan suatu inovasi ataupun suatu algoritma yang mampu mendeteksi COVID 19 dan dapat digunakan secara mandiri di rumah dan secara massal sehingga orang-orang yang menunjukkan gejala pernapasan COVID 19 tidak harus pergi ke rumah sakit, dengan begitu risiko adanya penularan virus juga dapat menurun.

Di era yang semakin maju dan berkembang ini, banyak penelitian dan juga teknologi baru yang dikembangkan untuk mengatasi suatu permasalahan di berbagai sektor, khususnya di sektor kesehatan. Salah satu perkembangan teknologi tersebut adalah dibidang kecerdasan buatan (*Artificial Intelligence*) yang saat ini

banyak digunakan untuk pendeteksian penyakit. Seperti yang saat ini sedang banyak dikembangkan, yaitu deteksi COVID 19.

Virus COVID 19 secara langsung menargetkan paru-paru dan mengakibatkan gejala pada saluran pernapasan seperti pneumonia dan sesak napas dengan konsekuensi dapat mengancam nyawa. Bagi penderita COVID 19 yang memiliki gejala utama seperti demam, batuk dan sesak napas akan memperlihatkan adanya gangguan pada sistem pernapasannya. Pemeriksaan terhadap suara napas dan pengaruhnya terhadap identifikasi komplikasi jalan napas dapat menunjukkan bahwa suara napas dari penderita COVID 19 memiliki pola sinyal akustik tertentu yang perlu diselidiki (Faezipour and Abuzneid, 2020). Suara napas tersebut kemudian dianalisis menggunakan pemrosesan dan analisis sinyal dan teknologi *artificial intelligence* (AI) seperti, *deep learning* dan *machine learning*, untuk menganalisis fase pernapasan pada paru-paru. Hal ini dapat digunakan sebagai algoritma pemantauan pernapasan mandiri untuk Pandemi COVID 19 global yang sedang berlangsung.

Teknologi kecerdasan buatan atau AI telah dimanfaatkan oleh para peneliti untuk mengembangkan metode deteksi dini COVID 19 dari sampel suara napas. Penelitian yang dilakukan oleh Sait tahun 2021 membandingkan diagnosis yang dilakukan untuk deteksi COVID 19 dan non COVID 19 menggunakan citra X-ray dan sampel suara napas. Metode yang diimplementasikan dalam penelitian tersebut adalah metode *Convolutional Neural Network* (CNN), yang merupakan salah satu metode *deep learning*. Hasil dari penelitian tersebut menunjukkan bahwa deteksi COVID 19 dengan citra X-ray mencapai tingkat akurasi 99,66%, sedangkan deteksi COVID 19 dengan sampel suara napas mencapai tingkat akurasi 80% (Sait et al., 2021).

Brown juga melakukan penelitian dengan sampel suara napas dan suara batuk untuk dapat mendeteksi COVID 19. Metode yang digunakan adalah *support vector machine* (SVM). Dalam penelitiannya tersebut digunakan total 74 data yang terdiri dari 54 data Covid 19 dan 20 data Non-Covid 19. Brown, mampu menunjukkan bahwa sampel suara napas dapat digunakan untuk deteksi kasus COVID 19 dengan nilai akurasi sebesar 80% (Brown et al., 2020). Kelemahan dari penelitian yang dilakukan oleh brown ini adalah data yang digunakan terbatas

sehingga interval kepercayaan relatif lebar.

Penggunaan metode deep learning yaitu, *Convolutional Neural Network* (CNN) dalam penelitian Sait memiliki kelemahan dalam proses pelatihannya. Pelatihan pada CNN membutuhkan waktu yang relatif lama dan sulit. Hal tersebut karena pelatihan pada CNN membutuhkan dataset yang besar dalam membangun model algoritma hingga siap digunakan. Selain itu, beban komputasi yang besar, menyebabkan spesifikasi *hardware* yang dibutuhkan menjadi tinggi (Nugroho et al., 2020).

*Machine learning* merupakan salah satu contoh dari pengembangan *artificial intelligence* (AI) yang memungkinkan sistem melakukan pembelajaran otomatis melalui suatu data pelatihan dengan mengidentifikasi suatu pola tertentu. Tidak seperti *deep learning*, *machine learning* memiliki model yang lebih sederhana sehingga tidak membutuhkan waktu yang lama dalam proses pelatihan model algoritmanya (Janiesch et al., 2021). Terdapat beragam algoritma *machine learning* yang biasa digunakan untuk klasifikasi, salah satunya adalah algoritma *k-Nearest Neighbor* (k-NN) yang merupakan algoritma sederhana yang menggunakan kesamaan ciri tetangga terdekat sebagai dasar dalam pengambilan keputusan (Singh and Pandey, 2016). k-NN memiliki kelebihan jika dibandingkan metode klasifikasi yang lain dimana metode tersebut tidak membutuhkan tahapan pelatihan terlebih dahulu dalam menjalankan algoritmanya, sehingga memungkinkan algoritma melakukan pembelajaran yang adaptif. Dalam penelitian yang dilakukan oleh Neili tahun 2020, diketahui bahwa metode k-NN mampu mengklasifikasikan suara napas normal dan suara napas abnormal dengan tingkat akurasi mencapai 95% menggunakan nilai k=10 pada fold data ke-10 (Neili et al., 2020). Don melakukan klasifikasi suara napas normal dan abnormal menggunakan metode k-NN sebagai pengklasifikasi dan didapatkan hasil akurasi sebesar 94,1% dengan menggunakan 5 fold cross validasi (Don, 2020). Kinerja dari metode k-NN ditentukan oleh jumlah data terdekat atau yang disebut dengan nilai k. Nilai k ini akan berpengaruh terhadap hasil akurasi yang akan diperoleh nantinya. Namun, penentuan jumlah nilai k tersebut merupakan kekurangan dari penerapan metode k-NN ini karena masih dilakukan secara manual dengan trial and error secara berulang, sehingga akan cukup sulit dalam memperoleh hasil klasifikasi dengan

nilai akurasi yang optimal.

Penambahan metode optimasi adalah salah satu cara yang dilakukan peneliti dalam mengatasi permasalahan yang terjadi pada proses klasifikasi untuk memperoleh hasil akurasi yang optimal. Beberapa metode optimasi yang telah dikembangkan terinspirasi dari berbagai peristiwa biologis di alam, salah satunya yaitu menggunakan pendekatan terhadap perilaku sekelompok hewan dalam mencari dan mengumpulkan makanan mereka. Salah satu contohnya adalah metode *Artificial Bee Colony* (ABC) yaitu, algoritma yang terinspirasi dari perilaku lebah madu dalam proses mencari sumber makan yang terbaik (Karaboga, 2005). Metode ABC ini memiliki kelebihan yaitu, algoritmanya sangat sederhana dan fleksibel jika dibandingkan dengan algoritma optimasi yang lain (Izzah, Dewi and Mutrofin, 2013). Pada penelitian sebelumnya Ariani telah menggunakan metode *Artificial Bee Colony* (ABC) sebagai metode optimasi yang digunakan untuk mengoptimasi nilai  $k$  pada  $k$ -NN. Dalam penelitian tersebut dijelaskan bahwa tingkat rata-rata akurasi hasil klasifikasi penyakit meningkat menjadi 98,72% dengan menggunakan optimasi ABC jika dibandingkan dengan metode  $k$ -NN saja (Ariani, 2019).

Selain nilai  $k$ , penggunaan metode ekstraksi fitur yang tepat dapat meningkatkan kemampuan suatu program dalam mengidentifikasi suatu kasus (Sankoh et al., 2015). Dengan menggunakan ekstraksi fitur, maka akan diperoleh suatu ciri khusus atau informasi spesifik, sehingga algoritma klasifikasi dapat melakukan prediksi yang baik dan tepat. Pada beberapa penelitian, banyak digunakan ekstraksi fitur dengan menggunakan metode analisis time domain dalam deteksi COVID 19 yang berbasis audio. Dragomir menggunakan ekstraksi fitur *approximate entropy* (ApEn) pada sinyal suara yang dihasilkan oleh suara jantung untuk mendeteksi *coronary occlusions*. Fitur ApEn ini merepresentasikan keteraturan pola untuk sinyal suara jantung tersebut. Didapatkan hasil 82% untuk sensitivitas sistem yang digunakan, yang mana dapat diketahui bahwa penggunaan fitur ApEn sebagai parameter deteksi dinilai sudah cukup baik. Sehingga fitur tersebut dapat di aplikasikan pada penelitian lain yang serupa. Selain ApEn, analisis fitur time domain berupa fitur statistik juga digunakan pada beberapa penelitian. Mondal dalam penelitiannya menggunakan fitur statistik untuk menganalisis suara paru-paru dan didapatkan tingkat akurasi 94.16%, sensitivitas 100% dan spesifisitas

93.75%.

Berdasarkan referensi jurnal penelitian sebelumnya maka akan dijadikan dasar untuk melakukan penelitian deteksi COVID 19 menggunakan suara napas. Kemudian untuk mendapatkan hasil ekstraksi fitur pada sinyal suara napas yang akan di klasifikasikan tersebut digunakan metode time domain analysis. Data hasil ekstraksi fitur ini akan dikelompokkan dalam 2 kelas yaitu COVID 19 dan non-COVID 19 (*healthy*) berdasarkan parameter fitur ekstraksi. Metode yang akan digunakan adalah *k-Nearest Neighbor* yang di optimasi dengan *Artificial bee colony*. Hasil yang diinginkan pada penelitian ini yaitu algoritma dapat mendeteksi data suara napas tergolong dalam kelas COVID 19 atau non COVID 19 (*healthy*) berdasarkan parameter ekstraksi fitur dengan tingkat akurasi, sensitivitas, dan spesifitas program yang baik. Akurasi sendiri merupakan tingkat kebenaran algoritma dalam mendeteksi data antara data uji dan data latih dari kelas COVID 19 dan non-COVID 19 (*healthy*). Sensitivitas merupakan hasil deteksi benar dari kelas COVID 19 yang dilakukan algoritma, sedangkan spesifisitas adalah hasil deteksi benar dari kelas -COVID 19 (*healthy*).

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah yang dapat disusun dalam penelitian ini adalah sebagai berikut :

1. Berapa jumlah populasi (SN) pada metode *artificial bee colony* dan nilai k pada metode *k-Nearest Neighbor* (k-NN) untuk menghasilkan kinerja sistem yang optimal?
2. Berapa tingkat akurasi, sensitivitas dan spesifisitas tertinggi untuk deteksi sinyal suara napas COVID 19?

## 1.3 Batasan Masalah

Batasan masalah di dalam penelitian ini dibuat agar tidak terjadi adanya perluasan topik masalah dalam penyajian penelitian. Adapun batasan masalah pokok tersebut sebagai berikut :

1. Data penelitian merupakan data sekunder yang diambil dari dataset Coswara yang merupakan sampel suara napas yang dikumpulkan melalui crowd-sourcing di seluruh dunia menggunakan aplikasi situs penelitian

arXiv:2005.10548 dengan judul “Coswara – A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis” oleh Neeraj Sharma, Prashant Krishnan, Rohit Kumar, Shreyas Ramoji, Srikanth Raj Chetupalli, Nirmala R., Prasanta Kumar Ghosh, dan Sriram Ganapathy.

2. Data yang digunakan pada penelitian ini hanya data sampel suara napas deep (Breathing Deep).
3. Keterangan pada dataset berupa usia, jenis kelamin, asal dan penyakit bawaan di abaikan.
4. Penelitian menggunakan data suara napas yang di klasifikasikan menjadi dua kelas yaitu COVID-19 dan non-COVID-19 sebagai data latih dan data uji.
5. Deteksi pada penelitian ini lebih mengarah pada klasifikasi.

#### **1.4 Tujuan Penelitian**

Tujuan dari penelitian ini adalah sebagai berikut :

1. Mendapatkan jumlah populasi (SN) pada metode artificial bee colony dan nilai  $k$  pada metode *k-Nearest Neighbor* (k-NN) untuk menghasilkan kinerja sistem yang optimal.
2. Mendapatkan tingkat akurasi, sensitivitas dan spesifitas tertinggi dalam deteksi sinyal suara napas berdasarkan hasil proses *k-Nearest Neighbor* (k-NN) yang telah dioptimasi dengan algoritma *artificial bee colony* (ABC).

#### **1.5 Manfaat Penelitian**

Penelitian yang dilakukan ini diharapkan dapat memberikan manfaat sebagai berikut :

1. Sarana deteksi dini untuk diagnosis awal kasus COVID 19 untuk menghindari penularan akibat pergi ke rumah sakit.
2. Menambah penelitian mengenai metode untuk deteksi awal dari sinyal suara napas untuk kasus COVID 19.
3. Referensi bagi para peneliti untuk pengembangan penelitian di dunia medis dimasa mendatang.