

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

*Coronavirus diseases* atau yang lebih dikenal dengan Covid-19 merupakan virus jenis baru yang disebabkan oleh virus SARS-Cov-2 [1]. Berdasarkan data dari World Health Organization (WHO) ditemukan pertama kali pada akhir bulan Desember 2019, di Kota Wuhan, China [2]. Virus ini kemudian menyebar ke seluruh negara, yang kemudian membuat WHO menetapkan virus ini sebagai pandemik pada Maret 2020 [3]. Sampai saat ini (9 Februari 2022) jumlah kasus Covid-19 diseluruh dunia mencapai 396.558.014 kasus, dengan kasus total kematian berjumlah 5.745.032 jiwa [3]. Di Indonesia sendiri, kasus pertama Covid-19 adalah pada awal Maret 2020, yang diumumkan langsung oleh Presiden Joko Widodo dan Menteri Kesehatan (Menkes) pada saat itu, yaitu, dr. Terawan Agus Putranto. Sampai saat ini (9 Februari 2022) jumlah kasus Covid-19 di Indonesia mencapai 4.580.093 kasus, dengan kasus total kematian berjumlah 144.719 jiwa [3].

Menurut Gut et al penyebaran Covid-19 berlangsung dalam kurun waktu yang sangat cepat, yang tentunya dapat mengancam hidup tiap individu, yang tidak terbatas aspek geografi maupun usai. Penyebaran Covid-19 diduga dapat menyebar melalui cairan tubuh manusia, melalui udara, dan melalui sentuhan benda yang telah terkontaminasi dengan Covid-19 [4]. Covid-19 juga bisa bertahan selama total waktu 72 jam pada benda-benda yang berbahan dasar plastik, 24 jam pada benda-benda yang berbahan dasar kardus, 4 jam pada benda-benda yang berbahan dasar tembaga, dan 3 jam di udara [5]. Orang yang telah terinfeksi dengan Covid-19 akan mengalami gejala dan bahkan bisa sampai menyebabkan kematian.

Adapun gejala-gejala umum yang dialami oleh orang yang telah terinfeksi Covid-19 adalah demam, sesak napas, batuk, kelelahan, sakit kepala, nyeri otot, mual, diare, bersin dan sakit tenggorakan, serta kehilangan kemampuan untuk mengenali bau dan rasa. Ada juga gejala-gejala yang tidak umum yang dapat dialami oleh orang yang telah terinfeksi Covid-19, yaitu; ruam kulit, konjungtivitis, panas dingin, pusing, cepat marah, cemas, atau depresi, serta gangguan tidur [6]. Tetapi terdapat kendala yang menyebabkan orang yang telah terinfeksi Covid-19 terkadang sulit untuk menyadari bahwa dirinya telah terinfeksi, sebab gejala Covid-

19 memiliki kesamaan dengan gejala yang terjadi pada penyakit Flu biasa. Bahkan gejala Covid-19 yang beragam juga dan memiliki kesamaan dengan beberapa penyakit lain, sehingga membuat orang kesulitan untuk membedakan antara Covid-19, penyakit Flu, ataupun penyakit lainnya.

Selain meningkatkan kewaspadaan serta kesadaran diri dalam merasakan gejala-gejala Covid-19, solusi dari permasalahan tersebut adalah dengan mengimplementasikan *Machine Learning* (ML) dengan Algoritma *Random Forest*. ML digunakan untuk mengklasifikasikan individu yang terpapar Covid-19 berdasarkan gejala-gejala yang ada. ML adalah salah satu cabang dari *Artificial Intelligence* dengan metode pembelajaran *Supervised Learning*.

Random Forest (RF) adalah salah satu metode yang terdapat pada *Ensemble Learning*. *Ensemble Learning* adalah teknik yang menggabungkan banyak metode klasifikasi agar mendapatkan solusi terbaik dari permasalahan yang kompleks [7]. *Ensemble Learning*, terbagi menjadi beberapa metode, yaitu; *Bagging*, *Boosting*, dan *Stacking*. RF merupakan salah satu contoh algoritma yang menggunakan metode *Bagging*. RF merupakan algoritma pembelajaran mesin yang digunakan untuk menyelesaikan masalah regresi dan klasifikasi [7]. RF juga merupakan pengembangan dari algoritma *Decision Tree*, sebab pada algoritma *Decision Tree* terdapat kelemahan yaitu apabila data set yang dimiliki berjumlah sedikit akan tetapi memiliki banyak atribut, serta tingginya potensi terjadinya *overfitting* [8]. Adapun beberapa keuntungan menggunakan algoritma RF adalah tingkat akurasi algoritma RF lebih tinggi dibandingkan menggunakan algoritma *Decision Tree*, menangani permasalahan *overfitting* pada algoritma *Decision Tree*. *Overfitting* yaitu ketika model yang dibuat terlalu fokus pada *training* dari dataset tertentu, yang membuat model tersebut tidak bisa melakukan prediksi dengan tepat jika diberikan dataset lain yang serupa. Algoritma *Random Forest* juga menangani *missing value* dengan cara yang efektif, serta menghasilkan prediksi yang sesuai walaupun tanpa penyetelan *hyper-parameter* [7].

Penggunaan Random Forest di bidang kesehatan sudah diterapkan pada berbagai penelitian, yaitu; prediksi kesehatan pasien Covid-19 dengan tingkat akurasi mencapai 94% serta F1 score adalah 0.86 [9], prediksi kanker payudara [10] dengan tingkat akurasi mencapai 100%, prediksi hasil virologi diantara orang dewasa yang terkena HIV [11], klasifikasi gangguan sakit kepala dengan tingkat akurasi terbaik mencapai 99.56% [12]. Bahkan terdapat penelitian yang menggunakan *highly imbalanced data* untuk prediksi 8 kategori penyakit, performa AUC (*Area under the ROC Curve*) yang didapat mencapai 88.79% [13].

Sehingga diharapkan, penggunaan algoritma *Random Forest* dapat membantu mengklasifikasikan individu yang terpapar covid-19, terpapar penyakit lain, atau tidak terpapar penyakit sama sekali.

## 1.2 Rumusan Masalah

Berdasarkan Latar Belakang yang telah disampaikan, maka Rumusan Masalah adalah sebagai berikut:

1. Bagaimana menerapkan *Random Forest* untuk melakukan klasifikasi individu yang terpapar Covid-19, Omicron, Influenza dan Pneumonia berdasarkan gejala yang diberikan?
2. Berapa nilai *Precision*, *Recall*, *F1-Score*, dan *Accuracy* yang didapat dengan mengimplementasikan algoritma *Random Forest*?

## 1.3 Batasan Masalah

Adapun Batasan Masalah pada penelitian ini adalah sebagai berikut:

1. Pakar yang dipilih adalah Dr. Lendy Steven Frederick Polii, S.Ked., yaitu seorang Dokter yang melakukan praktek di Kota Manado.
2. Dataset yang digunakan merupakan data Covid-19 di Kota Manado.
3. Dataset didapatkan dengan cara interview dan rekam medis, dari RSUP (Rumah Sakit Umum Pusat) Prof. Dr. R. D. Kandou Manado dan RS Siloam Manado.
4. Dataset yang digunakan didapat dari data yang diisi oleh dokter Dr. Carissa Vania Tirajoh, S.Ked., yang melakukan praktek di Kota Manado.

## 1.4 Tujuan Penelitian

Berdasarkan Rumusan Masalah yang telah disampaikan, Tujuan Penelitian adalah sebagai berikut:

1. Mengimplementasikan algoritma *Random Forest* untuk melakukan klasifikasi individu yang terpapar Covid-19, Omicron, Influenza, dan Pneumonia berdasarkan gejala-gejala yang diberikan.

2. Mengukur nilai *Precision*, *Recall*, *F1-Score*, dan *Accuracy* yang dihasilkan dalam menggunakan algoritma Random Forest.

### 1.5 Manfaat Penelitian

Manfaat penelitian ini adalah agar supaya dapat membantu pendeteksian infeksi Covid-19 dan Omicron, serta penyakit Influenza, dan Pneumonia pada tiap individu, sehingga bisa mendapatkan penanganan secepatnya. Penelitian ini juga diharapkan bisa menjadi referensi untuk penelitian-penelitian dimasa mendatang.

### 1.6 Sistematika Penulisan

Sistematika penulisan laporan adalah sebagai berikut:

- Bab 1 PENDAHULUAN

Bab pertama menjelaskan mengenai latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

- Bab 2 LANDASAN TEORI

Bab kedua berisi teori-teori yang digunakan sebagai acuan pada penelitian berupa Covid-19 serta gejala, algoritma *Random Forest*, dan evaluasi model.

- Bab 3 METODOLOGI PENELITIAN

Bab ketiga berisi tahap-tahap yang dilakukan selama penelitian berlangsung. Metodologi penelitian yang dilakukan adalah Studi Literatur, Pengumpulan Data, Praproses Data, Pembagian Data, Pembuatan Model, Evaluasi Model, serta Implementasi Model Pada Aplikasi.

- Bab 4 HASIL DAN DISKUSI

Bab keempat berisi tentang Implementasi, Analisis, dan Evaluasi yang dibuat berdasarkan hasil pemrograman. Bab keempat juga berisi integrasi model terbaik kedalam aplikasi.

- Bab 5 SIMPULAN DAN SARAN

Bab keelima berisi tentang Simpulan dari seluruh penelitian yang sudah dilakukan, serta Saran yang sekiranya dapat digunakan untuk penelitian-penelitian kedepannya.