

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

COVID-19 adalah penyakit yang disebabkan oleh keturunan *coronavirus* baru, ‘CO’ diambil dari corona, ‘VI’ virus, dan ‘D’ disease (penyakit) (Bender, 2020). Penyakit ini pertama kali ditemukan di Wuhan pada Desember 2019, Ibukota Provinsi Hubei China, dan telah menyebar secara global di seluruh dunia sejak itu, mengakibatkan pandemi *coronavirus* 2019-2020. Organisasi Kesehatan Dunia (WHO) mendeklarasikan wabah *coronavirus* 2019-2020 sebagai Kesehatan Masyarakat Darurat Internasional pada 30 Januari 2020, dan pandemi pada 11 Maret 2020 (Supriatna, 2020).

Organisasi Kesehatan Dunia (WHO) memberi nama virus baru tersebut *Severe Acute Respiratory Syndrome Coronavirus-2* (SARS-CoV-2) dan resmi menetapkan penyakit tersebut dengan nama *Novel Coronavirus* pada manusia dengan sebutan *Coronavirus Diseases 2019* atau yang lebih dikenal dengan sebutan COVID-19. Ketika menyerang manusia, virus ini menyebabkan infeksi saluran pernafasan seperti flu, hingga penyakit yang lebih berat seperti *Middle East Respiratory Syndrome* (MERS) dan *Severe Acute Respiratory Syndrome* (SARS). Sampai saat ini sudah dipastikan 222 Negara yang telah terjangkit, dengan kasus konfirmasi 127.877.462 dengan jumlah kematian 2.796.561 (Nabila et al., 2021).

Setiap hari kasus penyebaran COVID-19 di Indonesia terus meningkat, masyarakat diminta melakukan *social distancing* untuk memutus mata rantai penyebaran COVID-19 yang tersebar di berbagai wilayah di Indonesia (Sindi et al., 2020). COVID-19 pertama dilaporkan masuk ke Indonesia pada 2 Maret 2020, dengan 2 kasus terkonfirmasi. Wilayah Indonesia yang sangat luas memungkinkan perlunya pengelompokan bagian-bagian berdasarkan wilayah di Indonesia (Dwitri et al., 2020). Menurut data dari Kementerian Kesehatan Republik Indonesia per tanggal 31 Maret 2021 jumlah yang terkonfirmasi positif COVID-19 yaitu 1.511.712 dengan jumlah kematian 40,858 dan sembuh 1.348.330. Negara Indonesia merupakan Negara dengan kasus konfirmasi tertinggi di ASEAN (Kemenkes, 2021).

Saat ini *pneumonia virus corona* (COVID-19) yang disebabkan olehnya telah menyebar di China dan bahkan dunia. Simulasi yang akurat dan prediksi yang efektif dengan model matematika pada tren perkembangan pandemi ini menjadi sangat penting untuk pencegahan

han dan pengendalian pandemi (Vadyala et al., 2021; Xiao et al., 2021; Yoo et al., 2020; Alves et al., 2021). Beberapa penelitian telah dilakukan untuk memprediksi kasus-kasus baru COVID-19 di Indonesia dengan menggunakan algoritma-algoritma yang sudah ada seperti *Decision Tree* dan Regresi Linear. Terdapat beberapa kasus pula yang telah dilakukan dalam algoritma ini seperti Klasifikasi Penyakit ISPA (Pramadhani & Tedy, 2014), Prediksi Penyakit Hipertensi Kehamilan (Wulandari, 2016) dan Klasifikasi Data Peserta Didik (Sutoyo, 2018). Dengan kasus-kasus tersebut, dapat dikatakan bahwa algoritma yang dipakai cukup populer dalam kalangan penelitian.

Algoritma *Decision Tree* dan Regresi Linear sudah digunakan untuk berbagai klasifikasi maupun prediksi tetapi saat ini masih belum adanya perbandingan efektivitas terhadap dua algoritma ini dalam kasus prediksi penyebaran COVID-19 di Indonesia. Oleh karena itu, diperlukan suatu perbandingan antara nilai algoritma *Decision Tree* dan Regresi Linear untuk dapat menguji performa prediksinya dalam kasus prediksi penyebaran COVID-19 di Indonesia.

Berdasarkan latar belakang tersebut, maka peneliti tertarik untuk melakukan penelitian dengan judul “Perbandingan Nilai Algoritma *Decision Tree* dan Regresi Linear terhadap Kasus Prediksi Penyebaran COVID-19 di Indonesia”.

## 1.2 Rumusan Masalah

Berdasarkan permasalahan diatas, maka dapat dirumuskan beberapa masalah yaitu:

1. Bagaimana melakukan analisis terhadap data COVID-19 Indonesia untuk dapat mengetahui penyebaran COVID-19 di Indonesia?
2. Bagaimana melakukan prediksi terhadap data COVID-19 Indonesia untuk memprediksi kasus-kasus baru COVID-19 di Indonesia?
3. Bagaimana melakukan perbandingan nilai algoritma terhadap prediksi kasus penyebaran COVID-19 di Indonesia?

## 1.3 Batasan Masalah

Adapun batasan masalah dalam penelitian ini yaitu:

1. Hanya melakukan analisis dan prediksi COVID-19 Indonesia dari bulan Januari 2020 sampai Juli 2021.
2. Data yang diolah dalam penelitian menggunakan algoritma *Decision Tree* dan Regresi Linear.
3. Penerapan data mining ini menggunakan bahasa pemrograman Python.
4. Penelitian ini menggunakan penerapan *machine learning*.

#### **1.4 Tujuan Penelitian**

Adapun tujuan dari penelitian ini dilakukan berdasarkan rumusan masalah diatas yaitu:

1. Untuk mengetahui penyebaran COVID-19 di Indonesia dari hasil analisis data COVID-19 di Indonesia.
2. Membuat suatu prediksi dengan algoritma *Decision Tree* dan Regresi Linear terhadap kasus-kasus penyebaran COVID-19 di Indonesia.
3. Untuk dapat melakukan suatu perbandingan nilai algoritma *Decision Tree* dan Regresi Linear terhadap prediksi kasus penyebaran COVID-19 di Indonesia.

#### **1.5 Manfaat Penelitian**

Adapun manfaat dari penelitian yang dilakukan yaitu:

1. Sebagai alat bantu untuk memprediksi kasus-kasus baru COVID-19 di Indonesia.
2. Sebagai referensi mahasiswa lainnya untuk memberikan kontribusi dalam pengembangan isi penelitian ini.
3. Penelitian ini diharapkan dapat memberikan manfaat kepada pihak-pihak yang bersangkutan terkait hasil analisis dan prediksi COVID-19 di Indonesia.