

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pandemi COVID-19 yang disebabkan oleh virus corona dari subfamili Orthocoronavirinae, pertama kali muncul di kota Wuhan, Tiongkok pada akhir tahun 2019 dan dengan cepat menyebar ke berbagai negara, termasuk Indonesia pada Februari 2020. Virus ini menyebabkan infeksi saluran pernapasan yang dapat bervariasi dari gejala ringan seperti pilek hingga kondisi parah yang mengakibatkan pneumonia dan gagal organ. Pemerintah Indonesia, seperti banyak negara lain, menetapkan berbagai langkah pencegahan termasuk Pembatasan Sosial Berskala Besar (PSBB) dan penggolongan zonasi daerah untuk memantau dan mengendalikan penyebaran virus. Pemerintah menggunakan sistem zonasi berdasarkan risiko penyebaran virus, dengan zona hijau, kuning, oranye, dan merah untuk menggambarkan tingkat keamanan masing-masing daerah dengan adanya pembagian zona ini masyarakat akan lebih mudah untuk mengidentifikasi daerah-daerah yang rawan terpapar virus, hal ini sekaligus akan membantu mengurangi jumlah korban yang akan terpapar.

Perkembangan teknologi dan kecerdasan buatan (AI) menawarkan solusi untuk mengatasi keterbatasan akses informasi ini. Data mining, sebagai salah satu teknik dalam AI, dapat digunakan untuk menganalisis data penyebaran COVID-19 dan menghasilkan informasi yang berguna secara otomatis. Algoritma seperti K-Means dan Naive Bayes dapat membantu dalam mengelompokkan data ke dalam zona risiko dan melakukan klasifikasi yang akurat berdasarkan pola yang ditemukan dalam data tersebut.

Naïve Bayes Classifier (NBC) merupakan algoritma sederhana yang mampu melakukan klasifikasi dan merupakan sebuah pengelompokan berdasarkan pada teorema bayes. Metode *naïve bayes classifier* memiliki akurasi lebih baik dibandingkan dengan model klasifikasi *Neural Network* dan *Decision Tree*. *Naïve Bayes Classifier* mampu melakukan proses klasifikasi dan proses klasifikasi akan

lebih akurat jika data latih yang digunakan dalam training mempunyai jumlah yang lebih banyak. (Nurdin et al., 2021).

Penelitian sebelumnya telah menunjukkan bahwa algoritma Naive Bayes dapat memberikan akurasi yang baik dalam klasifikasi data COVID-19. Misalnya penelitian terdahulu yang telah dilakukan oleh (Risal et al., 2021) yang juga menjadi bahan referensi penelitian ini dengan judul “Penerapan Data Mining dalam Mengklasifikasikan Tingkat Kasus Covid-19 di Sulawesi Selatan Menggunakan Algoritma *Naïve Bayes*” pada penelitian ini 2 dari 23 data training tidak di klasifikasikan secara tepat sehingga penelitian ini mendapatkan tingkat akurasi sebesar 91%.

Penelitian lain yang berjudul “Penerapan Data Mining Untuk Klasifikasi Pandemi Covid-19 Di Indonesia Dengan Algoritma *Naïve Bayes*” oleh (Ermanto & Damas, 2020) dalam penelitian ini data yang telah disiapkan untuk klasifikasi dibagi menjadi 2 dengan menggunakan teknik *sampling random* data training (70%) dan data testing (30%) penelitian ini menghasilkan nilai *Precision* 100.00%, *Recall* 100.00% dan *Accuracy* 100.00%.

Penelitian ini bertujuan untuk mengembangkan aplikasi berbasis web yang memanfaatkan algoritma KMeans dan Naive Bayes untuk mengelompokkan dan mengklasifikasikan data penyebaran COVID-19 di Kota Lhokseumawe. Aplikasi ini tidak hanya memudahkan otoritas kesehatan dalam memantau dan mengendalikan penyebaran virus tetapi juga memberikan akses informasi yang lebih cepat dan mudah bagi masyarakat. Dengan demikian, masyarakat dapat lebih proaktif dalam mengambil tindakan pencegahan yang sesuai dengan kondisi zonasi di daerah mereka

Berdasarkan penelitian terdahulu, latar belakang dan belum adanya penelitian yang dilakukan sebelumnya di Lhokseumawe tentang klasifikasi daerah penyebaran COVID-19 dengan metode *Naïve Bayes* penulis memutuskan mengambil judul tugas akhir “Penerapan Metode *Naïve Bayes* Dalam Klasifikasi Daerah Penyebaran COVID-19 di Kota Lhokseumawe”.

1.2 Rumusan Masalah

1. Bagaimana caranya memproses dan mengelola data penyebaran COVID-19 di Kota Lhokseumawe secara efisien.?
2. Bagaimana cara menerapkan algoritma Naive Bayes untuk melakukan klasifikasi terhadap zona risiko penyebaran COVID-19 di Kota Lhokseumawe?
3. Seberapa akurat model Naive Bayes dalam memprediksi zona risiko penyebaran COVID-19 berdasarkan data yang telah diklusterkan?

1.3 Batasan Masalah

1. Data yang digunakan dari Dinas Kesehatan Kota Lhokseumawe.
2. Data yang telah diperoleh akan diolah dan dikelompokkan ke dalam 3 kelas (Zona Hijau, Zona Kuning dan Zona Merah) menggunakan metode K-Means karena data yang diperoleh tidak memiliki kelas.
3. Skalabilitas: Aplikasi ini dirancang untuk Kota Lhokseumawe, sehingga mungkin memerlukan penyesuaian lebih lanjut untuk digunakan di daerah lain dengan karakteristik yang berbeda.
4. Algoritma Terbatas: Penelitian ini menggunakan algoritma KMeans dan Naive Bayes. Algoritma lain mungkin memberikan hasil lebih baik tetapi tidak diimplementasikan dalam penelitian ini.

1.4 Manfaat Penelitian

Penulis berharap manfaat dari penelitian tugas akhir ini adalah sebagai berikut:

1. **Bantuan Pengambilan Keputusan:** Aplikasi ini dapat membantu pemerintah dan otoritas kesehatan dalam membuat keputusan berbasis data untuk mengelola penyebaran COVID-19 di Kota Lhokseumawe dengan lebih efektif.
2. **Penyajian Data yang Mudah Dipahami:** Dengan adanya visualisasi data dan hasil klustering, informasi mengenai penyebaran COVID-19 dapat disajikan dengan cara yang lebih mudah dipahami oleh masyarakat.
3. **Automasi Analisis Data:** Aplikasi ini mengotomatisasi proses analisis data, sehingga mengurangi beban kerja manual dalam memproses dan menganalisis data penyebaran COVID-19.

1.5 Tujuan Penelitian

1. Membangun aplikasi berbasis web yang dapat mengelola dan memproses data penyebaran COVID-19 di Kota Lhokseumawe secara efisien.
2. Mengimplementasikan algoritma KMeans untuk mengelompokkan data penyebaran COVID-19 ke dalam tiga zona risiko (hijau, kuning, merah).
3. Menerapkan algoritma Naive Bayes untuk melakukan klasifikasi zona risiko penyebaran COVID-19 berdasarkan data yang telah diklusterkan.