

Prediksi Status Darurat Covid-19 di Yogyakarta Menggunakan Naive Bayes

Prediction of the Covid-19 Emergency Status in Yogyakarta Using Naive Bayes

Tikaridha Hardiai*¹, Salsyabila Vidia Nur Afni²

^{1,2}Jurusan Teknologi Informasi, FST Universitas 'Aisyiyah Yogyakarta
e-mail: *1tikaridha@unisayogya.ac.id, 2salsyabila80@gmail.com

Abstrak

Meluasnya wabah virus corona di awal tahun 2020 menggemparkan dunia. Mudah-mudahan penyebaran virus tersebut, orang terkonfirmasi positif Covid-19 dengan kasus terkonfirmasi meluas terutama di Daerah Istimewa Yogyakarta (DIY). Tiap kabupaten di Daerah Istimewa Yogyakarta memiliki status tingkat gawat darurat Covid-19 yang berbeda-beda. Hal ini dapat berpengaruh pada status DIY berdasarkan kondisi kasus terkonfirmasi Covid-19. Pada penelitian ini menerapkan prediksi dari ilmu data mining untuk mengklasifikasi status kegawatdaruratan Covid-19 di DIY. Pengklasifikasian menggunakan metode Naive Bayes yang diterapkan untuk membantun model berdasarkan dataset pasien yang terkonfirmasi Covid-19. Model yang dibangun dapat memprediksi status gawat darurat DIY berdasarkan banyaknya kasus terkonfirmasi yang berada di tiap kabupaten di DIY. Pengolahan data diaplikasikan menggunakan Rapidminer. Hasil penelitian menunjukkan bahwa nilai akurasi mencapai 98,84%.

Kata kunci—Covid-19, Naive Bayes, prediksi, data mining, DIY

Abstract

The spread of the corona virus outbreak in early 2020 shocked the world. The easy spread of the virus, people who have been confirmed positive for Covid-19 with confirmed cases are widespread, especially in the Special Region of Yogyakarta (DIY). Each district in the Special Region of Yogyakarta has a different Covid-19 emergency level status. This can affect the status of DIY based on the condition of a confirmed case of Covid-19. In this study applied predictions from data mining science to classify the emergency status of Covid-19 in DIY. The classification uses the Naive Bayes method which is applied to help models based on a dataset of confirmed Covid-19 patients. The model built can predict DIY's emergency status based on the number of confirmed cases in each district in DIY. Data processing is applied using Rapidminer. The results showed that the accuracy value reached 98.84%.

Keywords—3-5 keywords, Algorithm A, B algorithms, complexity

1. PENDAHULUAN

Corona Virus Disease 2019 (Covid-19) adalah penyakit yang pertama kali muncul pada tahun 2019 yang menyebabkan peradangan paru-paru dan gangguan pernafasan [1]. Awal kemunculan Covid-19 diperkirakan berupa penyakit pneumonia, biasanya disertai gejala mirip

flu [1]. Gejala-gejala ini termasuk batuk, demam, kelelahan, sesak nafas dan kehi;angan nafsu makan. Karena penyebaran virus corona yang cepat, Organisasi Kesehatan Dunia (WHO) menyatakan virus corona sebagai pandemi pada tanggal 11 Maret 2020. Status pandemi menunjukkan fakta, Covid-19 menyebar begitu cepat sehingga hampir tidak ada negara di dunia yang dapat menjamin bahwa mereka terlindungi dari virus corona [1].

Setelah virus ini masuk ke Indonesia, menyebar dengan cepat di seluruh penjuru daerah di Indoensia, khususnya di Daerah Istimewa Yogyakarta. Banyaknya pasien di daerah Yogyakarta yang terinfeksi virus Covid-19 menjadi permasalahan utama yang dihadapi. Daerah yang menglam itingkat kegawat darurat tinggi memerlukan prioritas penanganan dibanding dengan daerah dengan tingkat kasus rendah.

Banyaknya pasien yang terkonfirmasi Covid-19 di DIY tiap-tiap Kabupaten menajdi permaslaha utama yang dihadapi pada setiap daerah Kabupaten. Daerah yang memiliki status kegawat darurat tinggi memiliki prioritas dalam penanganan dibanding daerah dengan status kegawat daruratan rendah atau sedang. Daerah dengan status gawat darurat tinggi memerlukan bantuan untuk mengklasifikasikan status daerahnya berdasarkan data pasien terkonfirmasi secara otomatis untuk meminimalisir kesalahan data terhadap kondisi status daerah. Oleh karena itu, diperlukan solusi teknologi otomatis berbasis data yang membantu mengkategorikan status gawat darurat di setiap Kabupaten. Peringkat kedaruratan masing-masing daerah dapat menjadi acuan Pemerintah untuk menentukan status level PPKM. Pengecekan status terkonfirmasi Covid-19 dan peningkatan mobilitas masyarakat DIY.

Berdasarkan data yang ada, tingkat penyebaran Covid-19 yang dikelompokkan tiap-tiap kab. Di DIY yang trinfeksi Covid-19. Klasifikasi ini merupakan cara yang tepat untuk mengetahui tingkat kedaruratan di DIY. Salah satunya dengan menerapkan Algoritma Naive Bayes untuk klasifikasi tingkat kegawat daruratan Covid-19 yang ada di DIY. Nantinya algoritma dapat digunakan untuk menyelesaikan masalah yang akan klasifikasi. Dari permasalahan diatas, dibutuhkan solusi teknologi yang dapat membantu melakukan klasifikasi status gawat darurat berdasarkan data pasien [2].

Data mining merupakan teknik untuk membuat pembelajaran machine learning menggunakan Teknik kecerdasan dengan protesis terkini yang belajar dalam membentuk contoh data empiris [3]. Data mining digunakan untuk menentukan pola dalam kumpulan besar data mentah [3]. Pada penelitian ini penulis menggunakan teknik data mining untuk mengklasifikasi dataset Covid-19 di tiap kabupaten menggunakan algoritma Naive Bayes.

Naive Bayes adalah salah satu algoritma yang digunakan dalam teknik klasifikasi data. *Naive Bayes* digunakan untuk teknik klasifikasi yang menggunakan metode probabilistik dan statistik untuk memprediksi kemungkinan masa depan berdasarkan pengalaman masa lalu, oleh karena itu dikenal dengan Teorema Bayes [4]. Menggunakan kombinasi *Naive Bayes*, dapat diasumsikan kondisi satu atribut dengan yang lain saling independen [5]. *Naive Bayes* mengasumsikan bahwa ada atau tidak adanya fitur tertentu dari satu kelas tidak ada hubungannya dengan fitur kelas lain.

Penelitian ini memprediksi status gawat darurat *Covid-19* di DIY menggunakan Teknik data mining berbasis dataset rata-rata pasien terkonfirmasi *Covid-19* tiap Kabupaten DIY. Metode yang dilakukan pada penelitian, menggunakan study literature, mengumpulkan dataset pasien terkonfirmasi tiap Kabupaten dan pembangunan Model *Naive Bayes*. Dalam pengolahan dalam pembangunan Model *Naive Bayes* menggunakan bantuan perangkat lunak yaitu *Rapidminer*.

Perbedaan penelitian yang dilakukan dengan penelitian sebelumnya terletak pada lokasi penelitian. Luaran penelitian yang dihasilkan berupa prediksi status kegawat darurat *Covid-19* di DIY, yang ditinjau berdasarkan hasil rerata kasus terkonfirmasi *Covid-19*.

Dalam menyusun penelitian ini, penulis banyak mereferensi dari penelitian-penelitian

sebelumnya yang ada kaitannya dengan penelitian ini. Dewi Yanti Liliana,dkk membahas prediksi status pasien *Covid-19* menggunakan *Naive Bayes*. Pada penelitian ini penulis mengambil dataset pasien *Covid-19* yang diperoleh dari www.kaggle.com. Hasil yang diperoleh memprediksi, berdasarkan usia dan jenis kelamin, kondisi pasien yang memiliki kemungkinan sembuh tertinggi dari *Covid-19*, dan pasien yang memiliki kemungkinan tinggi untuk tetap menjalani pengobatan dan dan meninggal. Hasil penelitian ini menunjukkan bahwa metode *Naive Bayes* memiliki tingkat akurasi yang tinggi dalam mengklasifikasikan status pasien yaitu sebesar 96,67% [6]. Selanjutnya penelitian yang dilakukan oleh Iqbal Try Chandra dan M. Husni Rifqo dengan judul penelitian "Penerapan Algoritma *Naive Bayes* dalam prediksi tingkat penyebaran *Covid-19* di Kota Bengkulu". Penelitian ini tentang prediksi tingkat penyebaran *Covid-19*. Data kriteria yang digunakan pada prediksi penyebaran *Covid-19*, data lokasi, suspek, positif, positif meninggal dan sembuh yang ada di Kota Bengkulu [7]. Luaran pada penelitian yang dilakukan Iqbal Try Chandra dan kawan, sistem yang dapat memberikan informasi prediksi keakurasian tingkat penyebaran *Covid-19* dengan menggunakan algoritma *Naive Bayes*, memprediksi tingkat keberhasilan algoritma *Naive Bayes* dalam memprediksi tingkat penyebaran *Covid-19* dan memperbudah pengguna dalam mengakses sistem dalam mencari informasi penyebaran *Covid-19*[7].

Penelitian ini bertujuan untuk memprediksi status kegawat darurat *Covid-19* di DIY berdasarkan rata-rata kasus *Covid-19* yang terkonfirmasi. Algoritma KNN dan *Naive Bayes* digunakan dalam penelitian, karena dirasa tepat dalam menghasilkan bentuk klasifikasi yang baik dan berguna untuk diimplementasikan secara nyata pada penanganan dan skrining kasus *Covid-19*.

Kontribusi dalam penelitian, untuk menghasilkan model yang memprediksi status darurat Covid 19 di DIY sehingga dapat mempersiapkan langkah-langkah penangan untuk menentukan status kabupaten pada zona merah, kuning atau hijau dalam jumlah kasus *Covid-19* yang terkonfirmasi.

2. METODE PENELITIAN

Metode yang digunakan pada penelitian menggunakan CRISP-DM (*Cross-Industry Standard Process for Data Mining*) dengan memberikan gambaran tentang siklus hidup proyek data mining. CRISP-DM merupakan standar yang telah dikembangkan pada tahun 1996 yang ditunjukkan untuk melakukan proses analisis dari suatu industri sebagai strategi pemecahan masalah dari bisnis atau unit penelitian. CRISP-DM menggunakan data mining yang sesuai untuk menyediakan proses standar umum dan non eksekutif untuk strategi pemecahan masalah bagi organisasi atau entitas penelitian.

Metode penelitian yang diusulkan peneliti meliputi beberapa tahapan dalam CRISP-DM yakni [8][9] [10]:

2.1 Pemahaman Bisnis (*Business Understanding*)

Tahap ini ditentukan tujuan penelitian kasus prediksi tingkat penyebaran *Covid-19*. Kemudian dilakukan penerjemahan dan analisis pada tujuan dan batasan yang menjadi formula pembatasan dari permasalahan data mining. Setelah itu dilakukan penyiapan strategi awal untuk mencapai tujuan.

Tahapan dalam pemahaman bisnis (*business understanding*) diantaranya:

- a. Menentukan tujuan proyek dan kebutuhan secara detail dalam lingkup bisnis atau unit penelitian secara keseluruhan.
- b. Menerjemahkan tujuan dan batasan menjadi formula dari permasalahan data mining.
- c. Menyiapkan strategi awal untuk mencapai tujuan.

2. 2 Pemahaman Data (*Data Understanding*)

Tahap ini dilakukan pengumpulan data untuk dianalisis dan diselidiki sehingga diketahui struktur pola data awal agar diperoleh gambaran pengetahuan awal yang akan dihasilkan dari data mining. Selanjutnya data tersebut dievaluasi kualitasnya untuk menghilangkan data-data missing value, duplikasi, dan tipografi. Jika memungkinkan dipilih sebagian kecil dari grup data yang mengandung pola dari permasalahan.

Tahapan dalam pemahaman data (*data understanding*) diantaranya:

- a. Mengumpulkan data, apabila dataset berasal dari lebih dari satu database, dilakukan proses integrasi data atau data integration.
- b. Mengembangkan analisis penyelidikan data untuk mengenali lebih lanjut data dan pencarian pengetahuan awal.
- c. Melakukan evaluasi kualitas data dengan memeriksa data dan membersihkan data yang tidak valid atau proses data cleaning.
- d. Memilih sebagian kecil grup data yang mengandung pola dari permasalahan.

2. 3 Pengolahan Data (*Data Preprocessing*)

Tahap ini dilakukan persiapan data dengan memilih kasus dan variabel yang ingin dianalisis dan yang sesuai dengan tipe analisis yang akan dilakukan. Kemudian dilakukan pemeriksaan pada data apakah perlu adanya perubahan pada beberapa variabel. Setelah tahap ini dilakukan diharapkan data siap dan memenuhi kriteria jika dilakukan pemodelan. Menambah data atau informasi lain yang dimungkinkan memudahkan proses data mining. Tahap ini memungkinkan untuk optimasi penentuan atribut, sehingga diperoleh atribut yang signifikan meningkatkan akurasi proses data mining.

Tahapan dalam pengolahan data (*data preprocessing*) diantaranya:

- a. Menyiapkan data awal, mengumpulkan data yang akan digunakan untuk keseluruhan tahapan berikutnya atau tahapan data selection.
- b. Memilih kasus dan variabel yang akan dianalisis sesuai dengan analisis yang dilakukan.
- c. Melakukan perubahan pada variabel apabila diperlukan.
- d. Menyiapkan data awal sehingga siap untuk perangkat pemodelan atau data transformation.

2. 4 Pemodelan (*Modeling*)

Pemodelan meliputi pengaturan situasi dan spesifikasi sehingga memungkinkan data dapat diproses menggunakan metode data mining yang direncanakan. Dalam tahap diperlukan tools atau pengkodean menggunakan bahasa pemrograman tertentu sehingga proses data mining dapat dikerjakan oleh sistem komputer. Pada tahap ini menggunakan software Rapidminer. Pemodelan menggunakan algoritma *Naive Bayes*.

Tahapan dalam pemodelan (*modeling*) diantaranya:

- a. Memilih dan mengaplikasikan teknik pemodelan yang sesuai.
- b. Melakukan kalibrasi aturan model untuk mengoptimalkan hasil.
- c. Dapat menggunakan beberapa teknik yang sama untuk permasalahan yang sama.
- d. Apabila terjadi kesalahan dalam proses modeling, kembali ke proses pengolahan data apabila diperlukan untuk membuat data dalam bentuk kebutuhan tertentu.

2.5 Tahap Evaluasi (*Evaluation*)

Mengevaluasi satu atau lebih model yang digunakan dalam pemodelan untuk mendapatkan kualitas dan efektifitas sebelum digunakan. Evaluasi dan validasi untuk performa (akurasi dan waktu proses) ketiga algoritma menggunakan confusion matrix dan kurva ROC. Dari hasil evaluasi dan validasi akan diketahui algoritma terbaik untuk menyelesaikan masalah.

Tahapan evaluasi (*evaluation*) diantaranya:

- a. Melakukan evaluasi satu atau lebih model yang digunakan, dalam penelitian menggunakan Model *Naive Bayes*.
- b. Menetapkan model yang diterapkan sudah sesuai dengan tujuan pada tahapan awal.
- c. Menentukan apakah terdapat permasalahan yang penting dari bisnis atau penelitian yang tidak tertangani dengan baik.
- d. Mengambil keputusan yang berkaitan dengan penggunaan hasil dari data mining.

2.6 Penyebaran (*Deployment*)

Penyebaran disini adalah penggunaan model yang dihasilkan yakni rangkaian metode dan data-data representatif yang sudah terolah sehingga memberikan informasi optimal ketika proses data mining. Penyebaran dalam lingkup sederhana adalah penggunaan hasil akhir dari data mining, saat mengklasifikasikan data menggunakan metode klasifikasi Naive Bayes.

3. HASIL DAN PEMBAHASAN

3.1 Pemahaman Bisnis

Pengimplementasian *data mining* pada penelitian ini berhubungan dengan data kasus terkonfirmasi Covid-19 di setiap Kabupaten di Daerah Istimewa Yogyakarta yang dijadikan acuan dalam langkah penanganan penentuan status zona setiap Kabupaten di DIY. Faktor-faktor yang mempengaruhi status zona diantaranya zona Hijau, Kuning, dan Merah setiap Kabupaten di DIY yaitu lonjakan rerata kasus terkonfirmasi pada setiap Kabupaten di DIY. Faktor-faktor tersebut mempengaruhi proses prediksi status zona sesuai dengan data yang sudah ada.

Berisi penjelasan tentang tahapan penelitian yang menggambarkan urutan logis untuk mendapatkan hasil penelitian sesuai dengan harapan dan gambaran penelitian. Jika ada gambar dan tabel, itu harus disajikan dengan nama tabel dan gambar yang disertai dengan nomor urut.

Tujuan pengimplementasian *data mining* agar mendapatkan pengetahuan tambahan mengenai penentuan prediksi status zona setiap Kabupaten di DIY agar dapat dilakukan proses skrining dampak Covid-19 di DIY.

Batasan-batasan dalam implementasi *data mining* pada penelitian ini diantaranya:

- a. Tidak ada faktor dari luar sumber dataset yang menjadi atribut yang dapat mempengaruhi proses implementasi. Seperti kasus pasien Covid-19 meninggal, kasus pasien Covid-19 sembuh.
- b. Jumlah dataset yang cukup banyak tidak semuanya pada dataset mempunyai data yang baik, sehingga dilakukan proses *data cleaning* untuk menghindari data missing atau data null. Agar memaksimalkan proses implementasi data sehingga pemrosesan lebih terfokus pada data pasien terkonfirmasi Covid-19.

Batasan diatas diharapkan dapat memaksimalkan data yang ada sehingga dapat mencapai hasil terbaik.

3.2 Pemahaman Data (*Data Understanding*)

Dataset rata-rata pasien Covid-19 tiap Kabupaten di DIY, akan dilantik untuk menghasilkan model prediksi terhadap status gawat darurat di Yogyakarta. Dataset yang diperoleh diambil mulai dari 19 Oktober 2021 – 13 Januari 2022. Dataset diperoleh dari data kasus terkonfirmasi positif Covid-19 dari website Dinas Kesehatan DIY. Dalam penentuan status zona didapat dari gabungan rata-rata tiap Kabupaten dengan kondisi yang ditunjukkan pada Tabel 1.

Tabel 1. Penentuan Status Zona

Kondisi	
Zona Merah	Rerata ≥ 7.5
Zona Kuning	Rerata ≤ 7.5
Zona Hijau	Rerata ≤ 4.00

Dinas Kesehatan DIY merilis status gawat darurat dengan klasifikasi zona resiko Covid-19, zona merah dengan range rata-rata ≥ 7.5 , zona kuning dengan range rata-rata ≤ 7.5 , zona hijau dengan range rata-rata ≤ 4.0 .

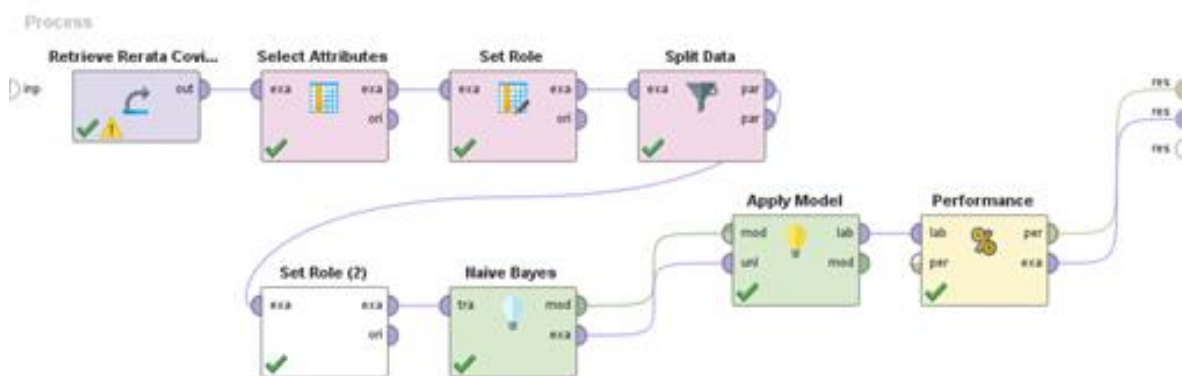
Data yang diambil selama rentang 19 Oktober 2021 – 13 Januari 2022, diperoleh pengumpulan sebanyak 144 data. Atribut yang terdapat pada dataset antara lain Tanggal, Sleman, Yogyakarta, Bantul, Kulon Progo, Gunung Kidul, Rerata, dan Status gawat darurat. Data yang digunakan adalah kasus terkonfirmasi setiap kabupaten di Yogyakarta. Sedangkan target kategori pasien yaitu rerata gabungan dari tiap kabupaten yang terkonfirmasi Covid-19.

3.3 Pengolahan Data

Pada tahapan preprocessing data, dilakukan tahap cleaning data terhadap data-data di tiap atribut karena memiliki data noise. Data yang diperoleh sebanyak 144, setelah dilakukan preprocessing data yang digunakan sebanyak 123 data. Setelah dilakukan preprocessing data, tahapan selanjutnya dilakukan pengujian data dengan menguji 86 data yang diaplikasikan dengan tools Rapidminer.

3.4 Pemodelan

Implementasi Data Mining dilakukan menggunakan perangkat lunak Rapidminer dengan Algoritma Naive Bayes. Naive Bayes Classifier adalah metode klasifikasi berdasarkan teorema Bayes. Metode pengklasifikasi dengan metode probabilitas dan statistik, yaitu dengan melakukan prediksi peluang masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Atribut yang terdapat dalam dataset diantaranya Kabupaten Sleman, Kabupaten Yogyakarta, Kabupaten Bantul, Kabupaten Kulon Progo, Kabupaten Gunung Kidul, rerata, dan status sebagai label atau target dari setiap kelas. Status dalam label atau target yang digunakan diantaranya status hijau, status kuning, dan status merah. Desain solusi Rapidminer menggunakan Naive Bayes ditunjukkan pada Gambar 1.

Gambar 1. Implementasi *Naive Bayes* dengan *Rapidminer*

Tujuan utama penelitian ini untuk memprediksi status gawat darurat setiap Kabupaten yang ada di DIY berdasarkan rata-rata kasus Covid-19 yang terkonfirmasi menggunakan algoritma Naive Bayes.

Hasil prediksi menggunakan Rapidminer sesuai dengan label data status tiap Kabupaten di Yogyakarta. Hasil prediksi klasifikasi Model Naïve Bayes menggunakan Rapidminer ditunjukkan pada Gambar 2.

Row No.	Status	predicti...	confide...	confide...	confide...	Sieman	Yogyak...	Bantul	Kulon P...	Gunung...	Rerata
1	Merah	Merah	1.000	0.000	0	19	8	3	6	25	12
2	Merah	Merah	1	0	0	41	20	4	31	15	22
3	Merah	Merah	1	0	0	36	13	5	11	4	13
4	Merah	Merah	1.000	0.000	0	3	11	6	15	18	10
5	Merah	Merah	0.981	0.019	0	5	11	6	12	6	8
6	Merah	Merah	1	0	0	36	10	3	18	17	16
7	Kuning	Kuning	0.218	0.782	0	7	5	5	2	15	6
8	Merah	Merah	1	0	0	28	10	3	17	17	15
9	Merah	Merah	1	0	0	24	15	4	18	6	13
10	Merah	Merah	0.714	0.286	0	22	0	5	13	1	8
11	Merah	Merah	1.000	0.000	0	39	4	3	15	8	13
12	Merah	Merah	1.000	0.000	0	21	3	2	2	19	9
13	Merah	Merah	1.000	0.000	0	19	4	2	24	2	10
14	Merah	Merah	1.000	0.000	0	25	14	2	2	5	9
15	Kuning	Kuning	0.123	0.877	0.000	19	5	2	5	5	7

ExampleSet (86 examples, 5 special attributes, 6 regular attributes)

Gambar 2. Hasil Prediksi dengan Rapidminer

3.5 Tahapan Evaluasi

Setelah dilakukan proses modeling menggunakan Model *Naive Bayes* dilakukan proses penilaian atau evaluasi terkait kualitas dan efektifitas model. Hasil evaluasi menggunakan Rapidminer sesuai dengan label asli data yang diuji pada pasien setiap Kabupaten di Yogyakarta.

Untuk menghitung performa dari klasifikasi dilakukan pengujian 86 data dari rentang tanggal 19 Oktober 2021 sampai dengan tanggal 13 Januari 2022. Rapidminer memiliki operator yang menampilkan hasil kinerja model dalam bentuk *Confussion Matrix*. Tabel 2 menunjukkan *Confussion Matrix* dari hasil pengujian.

Table 2. *Confussion Matrix* Hasil Pengujian

Accuracy: 98,84%				
	True Merah	True Kuning	True Hijau	Class Precision
Pred. Merah	19	0	0	100%
Pred. Kuning	1	24	0	96%
Pred. Hijau	0	0	42	100%
Class recall	95%	100%	100%	

Tabel *Confussion Matrix* menunjukkan informasi penting diantaranya hasil dari uji real dan jumlah data yang diuji dari hasil prediksi untuk setiap kelas yang diujikan. Berdasarkan *Confussion Matrix* hasil pengujian terlihat akurasi nilai Naive Bayes sebesar 98.84%. Nilai recall masing-masing kelas 95% untuk kelas Merah, 100% untuk kelas Kuning dan Hijau. Sedangkan nilai presisi masing-masing kelas sebesar 100% untuk kelas Merah dan Hijau dan 96% untuk kelas Kuning. Nilai akurasi ini disebabkan karena adanya beberapa data yang salah diklasifikasikan oleh sistem yaitu 1 data Merah diklasifikasikan menjadi data kelas Kuning.

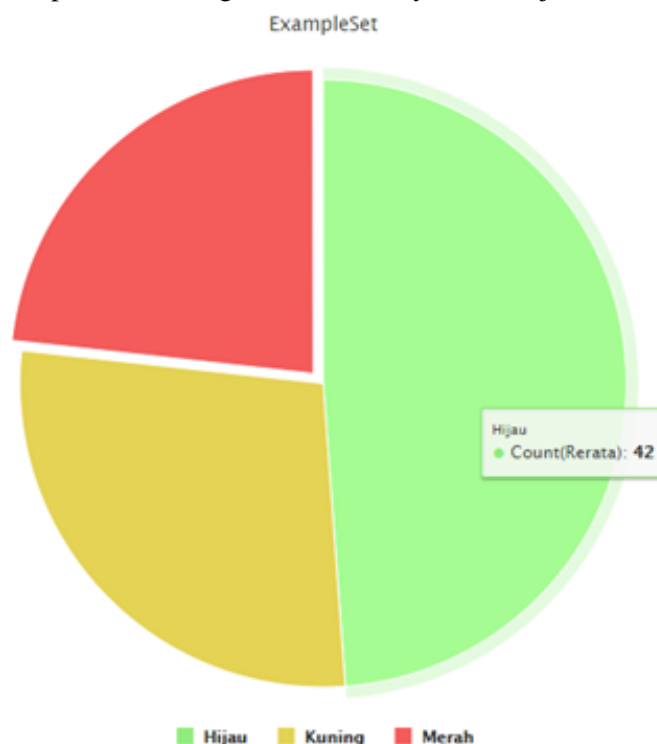
Confussion Matrix menunjukkan informasi penting diantaranya hasil uji rill dan jumlah data uji hasil prediksi untuk setiap kelas [11].

Nilai presisi menunjukkan seberapa akurat prediksi dalam mengidentifikasi data sesuai dengan kelas aslinya, sedangkan nilai recall menunjukkan kemampuan classifier untuk mengambil informasi sesuai dengan kelas aslinya. Akurasi menunjukkan kinerja semua model yang dibangun. Terlihat bahwa nilai akurasi klasifikasi dengan Naive Bayes mencapai 98,84%. Menunjukkan model yang dibangun memiliki sensitivitas yang baik dalam mengidentifikasi pengujian yang ada, sedangkan akurasi mengacu pada keakuratan hasil prediksi.

3.6 Tahapan Deployment

Setelah tahapan evaluasi dimana dilakukan proses penilaian secara detail terkait proses pemodelan menggunakan Model *Naive Bayes*. Pengimplementasian keseluruhan Model yang dirancang, selain itu dilakukan penyesuaian dari model dengan proses dari sistem Rapidminer sehingga menghasilkan hasil yang sesuai dengan target awal tahapan CRISP-DM.

Detail dari hasil pelaporan berupa grafik ditunjukkan pada Gambar 3. Proses mining terdapat label yang dilaporkan dalam grafik, diantaranya label Hijau, Kuning, dan label Merah.



Gambar 3. Detail Analisa *Naive Bayes*

Detail dari analisa menunjukkan bahwa kelas terbanyak tiap-tiap kabupaten menunjukkan kelas Hijau dengan rata-rata 42 yang berarti daerah tersebut termasuk dalam zona aman kasus terkonfirmasi Covid-19. Disusul kelas Kuning yang berarti daerah tersebut masuk ke dalam zona dengan kasus terkonfirmasi Covid-19 rendah, selanjutnya kelas Merah yang berarti daerah tersebut termasuk dalam zona tidak aman dengan kasus terkonfirmasi Covid-19 tinggi.

4. KESIMPULAN

Penelitian ini menghasilkan kesimpulan data yang dikembangkan dengan Model Data Mining untuk prediksi status gawat darurat di Yogyakarta menggunakan Model *Naive Bayes*. Hasil dari Model *Naive Bayes* yang diimplementasikan menggunakan perangkat lunak *Rapidminer* untuk memprediksi status gawat darurat DIY memberikan hasil akurasi sebesar 98,84%. Hasil penelitian ini bermanfaat untuk menerapkan pada situasi nyata untuk membantu Pemerintah Daerah Istimewa Yogyakarta menentukan status level PPKM atau darurat *Covid-19*. Jumlah dataset nyata dan berjumlah lebih besar dengan proporsi nilai setiap kelas yang seimbang sangat baik untuk mendapatkan akurasi prediksi yang tinggi.

5. SARAN

Penelitian ini jauh dari kesempurnaan, banyak kekurangan yang perlu penulis perbaiki. Hal ini dikarenakan masih minimnya data yang diperoleh penulis. Penelitian selanjutnya dapat menggunakan algoritma berbeda. Kritik dan saran membangun dari pembaca diharapkan sebagai bahan evaluasi untuk ke depannya.

DAFTAR PUSTAKA

- [1] Alvina Felicia Watratan, Arwini Puspita. B, and Dikwan Moeis, "Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran *Covid-19* Di Indonesia," *Journal of Applied Computer Science and Technology*, vol. 1, no. 1, pp. 7–14, 2020, doi: 10.52158/jacost.v1i1.9.
- [2] A. Hendriani and S. K. Sianturi, "Sentiment Masyarakat Terhadap Virus *COVID-19* Pada Instagram Menggunakan Algoritma Naive Bayes Classifier," *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, vol. 5, pp. 443–452, 2021.
- [3] F. Adiba, "Penerapan Data Mining dalam Mengklasifikasikan Tingkat Kasus *Covid-19* di Sulawesi Selatan Menggunakan Algoritma Naive Bayes," *Indonesian Journal of Fundamental Sciences*, vol. 7, no. 1, pp. 18–28, 2021.
- [4] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *ILKOM Jurnal Ilmiah*, vol. 10, no. 2, pp. 160–165, 2018, doi: 10.33096/ilkom.v10i2.303.160-165.
- [5] B. A. Mukhtar, N. A. Setiawan, and T. B. Adji, "Pembobotan Korelasi Pada Naive Bayes Classifier," *Seminar Nasional Teknologi Informasi dan Multimedia 2015*, no. 2, pp. 43–47, 2015.
- [6] D. Y. Liliana, H. Maulana, and A. Setiawan, "Data Mining untuk Prediksi Status Pasien *Covid-19* dengan Pengklasifikasi Naive Bayes," *Multinetics*, vol. 7, no. 1, pp. 48–53, 2021, doi: 10.32722/multinetics.v7i1.3786.
- [7] D. I. K. Bengkulu, "Penerapan Algoritma Naive Bayes Dalam Prediksi Tingkat Penyebaran Corona Virus Disease 19 (*Covid-19*)," Vol. 19, 2020.
- [8] T. Hardiani, "Analisis Clustering Kasus Covid 19 Di Indonesia Menggunakan Algoritma K-Means," Vol. 11, No. 2, Doi: 10.23887/Janapati.V11i2.45376.
- [9] T. Hardiani, "Comparison of Naive Bayes Method, K-NN (K-Nearest Neighbor) and Decision Tree for Predicting the Graduation of 'Aisyiyah University Students of Yogyakarta,'" *International Journal of Health Science and Technology*, vol. 2, no. 1, pp. 75–85, Jan. 2021, doi: 10.31101/ijhst.v2i1.1829.
- [10] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. 2012. doi: 10.1016/B978-0-12-381479-1.00001-0.
- [11] S. Hikmawan, A. Pardamean, and S. N. Khasanah, "Sentimen Analisis Publik Terhadap Joko Widodo terhadap wabah *Covid-19* menggunakan Metode Machine Learning," *J. Kaji. Ilm.*, vol. 20, no. 2, pp. 167–176, 2020, doi: 10.31599/jki.v20i2.117.

- [12] D. M. umi Atmaja, A. R. Hakim, D. Haryadi, and N. Suwaryo, “Penerapan Algoritma K-Nearest Neighbor Untuk Prediksi Pengelompokan Tingkat Risiko Penyebaran *COVID-19* Jawa Barat,” *Sntem*, vol. 1, no. November, pp. 1218–1226, 2021.
- [13] R. Ridho and H. Hendra, “Klasifikasi Diagnosis Penyakit *Covid-19* Menggunakan Metode Decision Tree,” *JUST IT J. Sist. Informasi, Teknol. Inf. dan Komput.*, vol. 11, no. 3, pp. 69–75, 2022, [Online]. Available: <https://jurnal.umj.ac.id/index.php/just-it/article/view/13594>.
- [14] Y. Purnamasari and Y. N. Kunang, “Pemanfaatan Data Mining Dalam Memprediksi Kasus Positif *Covid-19* Di Kota Palembang Menggunakan Algoritma K-Nearest Neighbors,” *J. Softw. Eng. Ampera*, vol. 2, no. 2, pp. 118–128, 2021, doi: 10.51519/journalsea.v2i2.128.
- [15] Rayuwati, Husna Gemasih, and Irma Nizar, “IMPLEMENTASI AIGORITMA NAIVE BAYES UNTUK MEMPREDIKSI TINGKAT PENYEBARAN COVID,” *Jural Ris. Rumpun Ilmu Tek.*, vol. 1, no. 1, pp. 38–46, 2022, doi: 10.55606/jurritek.v1i1.127.