

Implementasi Algoritma Naive Bayes Untuk Memprediksi zona dan Tingkat Penyebaran Covid-19 Di Provinsi Jambi

Jopi Mariyanto¹, Sandi Pramadi², Kholil Ikhsan³, Imelda Yose⁴, Errissya Rasywir^{5*}, Marrylinteri Istoningtyas⁶, Yovi Pratama⁷

¹Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Dinamika Bangsa, Kota Jambi, Indonesia

Email : ¹jopimariyanto2@gmail.com, ²sandipramadi.13@gmail.com, ³kholilikhshan@gmail.com,

⁴errissya.rasywir@gmail.com, ⁵imel_dyose@yahoo.com, ⁶marrylinteri_i@stikom-db.ac.id,

⁷yovi.pratama@gmail.com

Email Penulis Korespondensi : errissya.rasywir@gmail.com

Abstrak - Pada masa pandemi Virus Corona (Covid-19), Informasi merupakan hal yang sangat penting bagi masyarakat. Salah satu sumber informasi yang digunakan melalui internet yaitu website berita seperti kompas.com, detik.com, tribunnews.com. Namun, artikel berita terkait topik pembahasan kematian pasien, pasien yang sembuh dari Virus Corona dan zona pada suatu tempat belum dapat dikelompokkan. Hal ini dikarenakan angka pasien yang meninggal dan angka pasien yang sembuh akibat virus corona terus bertambah. Oleh karena itu, dibutuhkannya sistem yang dapat mengelompokkan berita berdasarkan data yang telah ada sebelumnya. Hal ini dapat diatasi menggunakan text mining, salah satunya metode klasifikasi pada text mining dapat mengelompokkan data pada suatu objek yang belum diketahui sebelumnya.

Kata Kunci : covid-19, klasifikasi, algoritma naive bayes

Abstract - During the Corona Virus (Covid-19) pandemic, information is very important for the community. One source of information used through the internet is news websites such as kompas.com, detik.com, tribunnews.com. However, news articles related to the topic of discussing patient deaths, patients recovering from the Corona Virus and zones in one place cannot be grouped. This is because the number of patients who died and the number of patients who recovered from the corona virus continued to increase. Therefore, we need a system that can classify news based on pre-existing data. This can be overcome using text mining, one of which classification methods in text mining can group data on an object that has not been previously known.

Keywords : covid-19, classification, nave bayes algorithm

1. PENDAHULUAN

Dunia saat ini tengah waspada dengan penyebaran sebuah virus yang dikenal dengan virus corona. Corona viruses (CoV) merupakan bagian dari keluarga virus yang menyebabkan penyakit mulai dari flu hingga penyakit yang lebih berat seperti Middle East Respiratory Syndrome (MERS-CoV) and Severe Acute Respiratory Syndrome (SARS-CoV) [1]–[3]. Penyakit yang disebabkan virus corona, atau dikenal dengan COVID19, adalah jenis baru yang ditemukan pada tahun 2019 dan belum pernah diidentifikasi menyerang manusia sebelumnya (World Health Organization, 2019). Kasus virus corona muncul dan menyerang manusia pertama kali di provinsi Wuhan, China. Awal kemunculannya diduga merupakan penyakit pneumonia, dengan gejala serupa sakit flu pada umumnya. Gejala tersebut di antaranya batuk, demam, letih, sesak napas dan tidak nafsu makan. Namun berbeda dengan influenza, virus corona dapat berkembang dengan cepat hingga mengakibatkan infeksi lebih parah dan gagal organ. Kondisi darurat ini terutama terjadi pada pasien dengan masalah kesehatan sebelumnya. Karena penularan virus corona yang sangat cepat inilah Organisasi Kesehatan Dunia (WHO) menetapkan virus corona sebagai pandemi pada 11 Maret 2020. Status pandemic atau epidemi global menandakan bahwa penyebaran COVID-19 berlangsung sangat cepat hingga hampir tak ada negara di dunia yang dapat memastikan diri terhindar dari virus corona[4]. Setelah virus ini masuk ke Indonesia, menyebarlah dengan cepat di seluruh penjuru daerah Indonesia termasuk Provinsi Jambi[5].

Berdasarkan data yang ada, tingkat penyebaran COVID-19 ini dapat dikelompokkan tiap-tiap daerah di Provinsi Jambi yang terinfeksi COVID-19 ini. Dalam hal ini klasifikasi merupakan cara yang tepat untuk mengetahui tingkat penyebaran COVID-19. Salah satunya adalah dengan menerapkan Algoritma Naive Bayes untuk klasifikasi tingkat penyebaran COVID-19 yang ada di provinsi Jambi. Klasifikasi merupakan salah satu poin utama dalam pengolahan data mining. Klasifikasi juga merupakan pengelompokan data dimana data yang digunakan mempunyai kelas atau target yang akan diklasifikasi. Sehingga nantinya algoritma dapat digunakan untuk menyelesaikan masalah yang akan klasifikasi [5].

Algoritma Naive Bayes [6]–[9] itu sendiri merupakan salah satu algoritma yang digunakan untuk teknik klasifikasi data. Naive Bayes digunakan untuk teknik pengklasifikasian data dengan metode kemungkinan dan statistik yang memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan Naive Bayes dimana diasumsikan kondisi antara atribut yang satu dengan atribut yang lainnya itu saling bebas. Klasifikasi Naive Bayes diasumsikan bahwa ada atau tidak adanya ciri-ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas yang lainnya [10].

Pada penelitian ini kami akan melakukan penerapan algoritma Naive Bayes untuk memprediksi tingkat penyebaran COVID-19 di provinsi Jambi dengan mengimplementasikan Algoritma Naive Bayes, yang berguna sebagai langkahantisipasi pandemic COVID-19.

2. METEDOLOGI PENELITIAN

2.1 Naive Bayes

Naive Bayes Classifier adalah metode klasifikasi atau teknik machine learning yang populer/umum digunakan dalam klasifikasi teks, memiliki performa yang baik pada banyak domain, sederhana dan efisien. Namun Naive Bayes sangat sensitive dalam pemilihan fitur, terlalu banyak jumlah fitur dapat mengakibatkan meningkatnya waktu perhitungan dan menurunkan akurasi klasifikasi [11].

2.2 Analisis Penerapan Naive Bayes

Untuk memperoleh suatu kesimpulan dari hasil pengumpulan data maka dilakukan analisis terhadap semua data yang terkumpul dan metode yang digunakan sesuai kebutuhan.

2.3 Pengumpulan Data

Pengumpulan data yang kami lakukan melalui situs <https://corona.jambiprov.go.id/> tentang tingkat penyebaran COVID-19 di provinsi Jambi dan dapat dikelompokkan tiap-tiap daerah di Provinsi Jambi yang terinfeksi COVID-19 dan juga memprediksi Zona di setiap daerah tersebut.

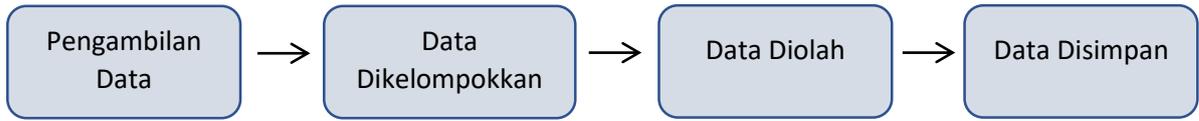
a. Studi Literatur

Studi literatur yang digunakan dalam penulisan Penelitian ini adalah mengumpulkan bahan referensi mengenai metode *klasifikasi* dari berbagai buku, jurnal, dan beberapa referensi lainnya.

b. Observasi

Metode pengumpulan data yang dilakukan dengan mengadakan pengamatan langsung terhadap sumber permasalahan.

2.4 Skenario Pengambilan Data Pada <https://corona.jambiprov.go.id/>



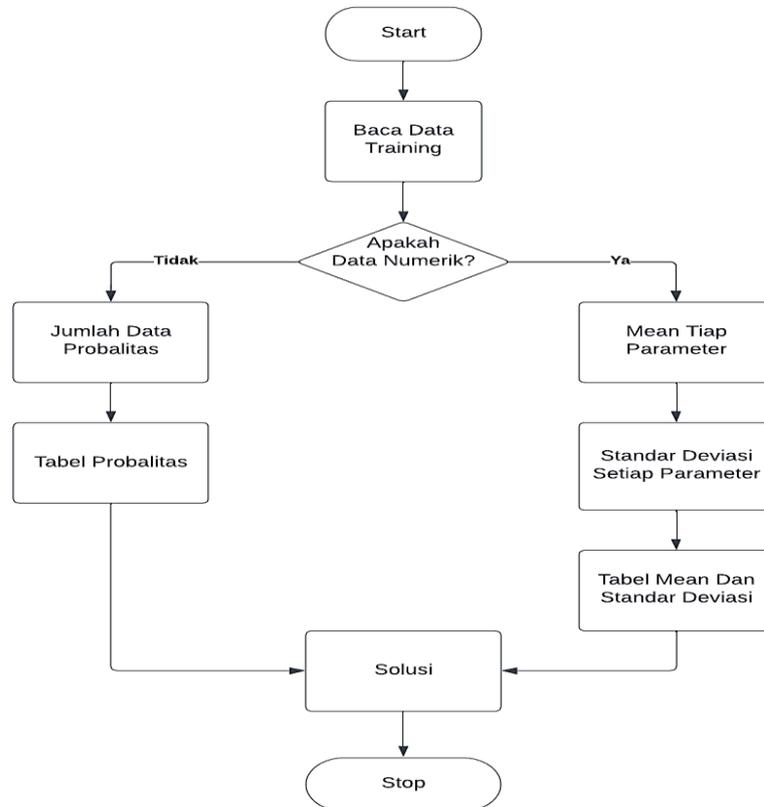
Skenario dalam pengambilan data:

1. Pengambilan data diambil yang diperoleh dari situs <https://corona.jambiprov.go.id/> .
2. Data dikelompokkan menjadi 5 atribut berdasarkan data aslinya.
3. Data diolah sedemikian rupa dengan meminimalkan/menghilangkan beberapa atribut.
4. Data disimpan dalam file berekstensi .xlsx

3. HASIL PEMBAHASAN

3.1 Implementasi Algoritma *Naive Bayes*

Pada bagian ini Ketika data siap di proses, algoritma *Naive Bayes* memproses data[12] pada gambar dibawah ini



Gambar 1. Tahapan Algoritma *Naive Bayes*

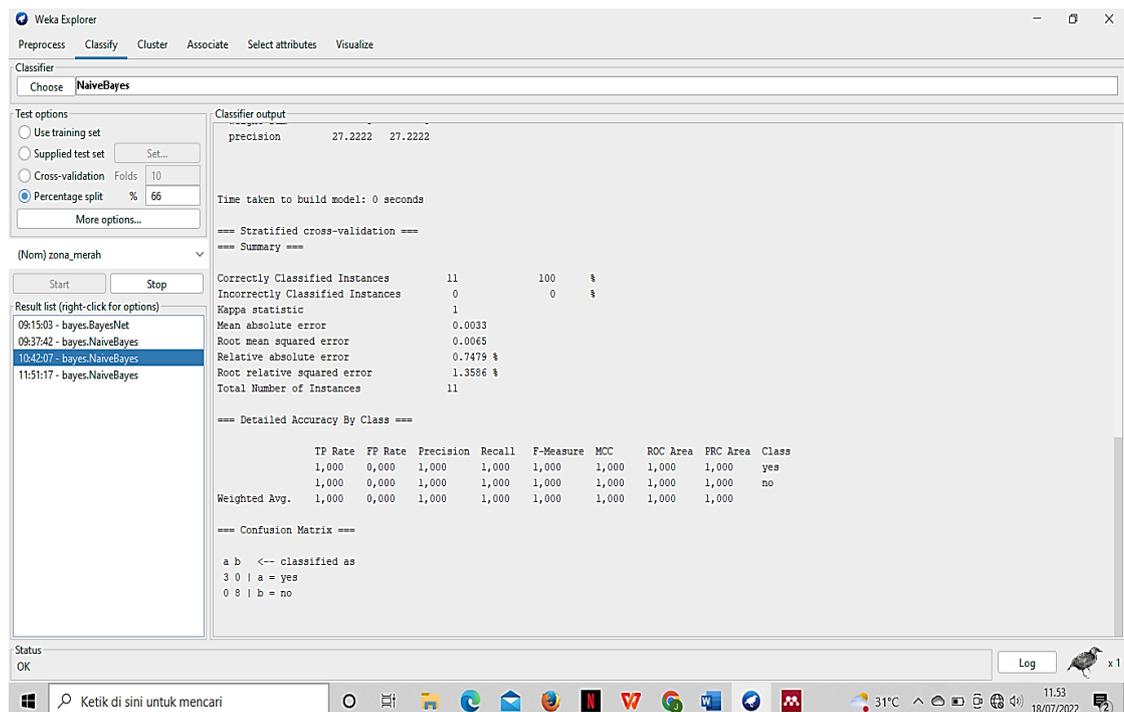
Tabel berikut merupakan data Covid-19 provinsi Jambi :

Tabel 1. Training Set Yang Telah Diolah

No	Kab/Kota	Positif	Sembuh	Meninggal	Zona Merah
1	batanghari	3165:00:00	3048:00:00	114:00:00	yes
2	bungo	2077:00:00	2046:00:00	31:00:00	no
3	kerinci	590:00:00	564:00:00	22:00	no
4	kota_jambi	9738:00:00	9482:00:00	254:00:00	yes
5	merangin	1679:00:00	1585:00:00	94:00:00	no
6	muaro_jambi	4131:00:00	4060:00:00	68:00:00	yes
7	sarolangun	1036:00:00	1027:00:00	09:00	no
8	sungai_penuh	1408:00:00	1388:00:00	20:00	no
9	tanjab_barat	2419:00:00	2353:00:00	63:00:00	no
10	tanjab_timur	1406:00:00	1370:00:00	36:00:00	no
11	tebo	2137:00:00	2067:00:00	31:00:00	no

3.2 Implementasi Klasifikasi Naive Bayes Pada Weka

Dengan Menggunakan *tools* Weka, data akan diuji seperti pada gambar 3 berikut :



Gambar 3. Klasifikasi Metode Naive Bayes

Berdasarkan Gambar 3 diatas dapat dilihat persentase untuk *correctly classified instance* adalah sebesar 100% sementara persentase untuk *Incorrectly Classified Instance* adalah sebesar 0%. Dimana 11 data Penyebaran Covid-19 semuanya berhasil diklasifikasikan dengan benar. Dapat disimpulkan bahwa *corectly clasified* yang didapatkan adalah 100% sedangkan untuk *Incorectly Clasified* yang didapatkan adalah 0%.

4. KESIMPULAN

Berdasarkan penelitian tentang Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid 19 di provinsi Jambi dapat ditarik kesimpulan yaitu Metode Naive Bayes memanfaatkan data training untuk menghasilkan probabilitas setiap kriteria untuk class yang berbeda, sehingga nilai-nilai probabilitas dari kriteria tersebut dapat dioptimalkan untuk memprediksi Tingkat Penyebaran Covid 19 di Provinsi Jambi berdasarkan proses klasifikasi yang dilakukan oleh metode Naive Bayes itu sendiri. Berdasarkan data Kasus COVID-19 Per Kab/Kota yang dijadikan data training, metode Naive Bayes berhasil mengklasifikasikan 11 yang diuji. Algoritma Naive Bayes di dukung oleh ilmu Probabilistik dan ilmu statistika khususnya dalam penggunaan data petunjuk untuk mendukung keputusan pengklasifikasian Pada algoritma Naive Bayes. semua atribut akan memberikan kontribusinya dalam pengambilan keputusan, dengan bobot atribut yang sama penting dan setiap atribut saling bebas satu sama lain.

5. SARAN

Adapun Saran yang diberikan mengenai penelitian tentang Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid 19 di Provinsi Jambi yaitu pengujian, sebaiknya dilakukan dengan menggunakan metode lain untuk melihat metode mana yang lebih akurat dalam memprediksi Tingkat Penyebaran Covid 19 di Provinsi Jambi.

REFERENCES

- [1] G. Aprilian Anarki, K. Auliasari, and M. Orisa, "PENERAPAN METODE HAAR CASCADE PADA APLIKASI DETEKSI MASKER," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 5, no. 1, pp. 179–186, 2021, doi: 10.36040/jati.v5i1.3214.
- [2] N. Dwitri, J. A. Tampubolon, S. Prayoga, F. Ilmi Zer, and D. Hartama, "Penerapan Algoritma K-Means Dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 Di Indonesia," *Jti (Jurnal Teknol. Informasi)*, vol. 4, no. 1, pp. 101–105, 2020.
- [3] A. R. Setiawan, "Lembar Kegiatan Literasi Sainifik untuk Pembelajaran Jarak Jauh Topik Penyakit Coronavirus 2019 (COVID-19)," *Edukatif J. Ilmu Pendidik.*, vol. 2, no. 1, pp. 28–37, 2020, doi: 10.31004/edukatif.v2i1.80.
- [4] Alvina Felicia Watratan, Arwini Puspita. B, and Dikwan Moeis, "Implementasi Algoritma Naive Bayes Untuk Memprediksi Tingkat Penyebaran Covid-19 Di Indonesia," *J. Appl. Comput. Sci. Technol.*, vol. 1, no. 1, pp. 7–14, 2020, doi: 10.52158/jacost.v1i1.9.
- [5] M. H. R. Ikkal Try Chandral, "PENERAPAN ALGORITMA NAÏVE BAYES DALAM PREDIKSI TINGKAT PENYEBARAN CORONA VIRUS DISEASE 19 (COVID-19) DI KOTA BENGKULU," p. 18, 2019.
- [6] H. Prastiwi, J. Pricilia, and E. Raswir, "Implementasi Data Mining Untuk Menentuksn Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM)," *J. Inform. Dan Rekayasa Komput.*, vol. 1, no. April, pp. 141–148, 2022.
- [7] M. R. Borroek, E. Rasywir, Y. Pratama, Fachruddin, and M. Istoningtyas, "Analysis on Knowledge Layer Application for Knowledge Based System," in *Proceedings of 2018 International Conference on Electrical Engineering and Computer Science, ICECOS 2018*, 2019, pp. 177–182, doi: 10.1109/ICECOS.2018.8605262.
- [8] Fachruddin, Saparudin, E. Rasywir, and Y. Pratama, "Network and layer experiment using convolutional neural network for content based image retrieval work," *Telkonnika (Telecommunication Comput. Electron. Control.*, vol. 20, no. 1, pp. 118–128, 2022, doi: 10.12928/TELKOMNIKA.v20i1.19759.
- [9] E. Rasywir, R. Sinaga, and Y. Pratama, "Analisis dan Implementasi Diagnosis Penyakit Sawit dengan Metode Convolutional Neural Network (CNN)," *J. Paradig. UBSI*, vol. 22, no. 2, pp. 117–123, 2020.
- [10] S. Syarli and A. Muin, "Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi)," *J. Ilm. Ilmu Komput.*, vol. 2, no. 1, pp. 22–26, 2016.

- [11] D. A. Muthia, "Opinion Mining Pada Review Buku Menggunakan Algoritma Naive Bayes," *J. Tek. Komput.*, vol. 2, no. 1, pp. 1–8, 2016.
- [12] D. Darwis, N. Siskawati, and Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *J. Tekno Kompak*, vol. 15, no. 1, p. 131, 2021, doi: 10.33365/jtk.v15i1.744.