

Penerapan Algoritma Naïve Bayes Dalam Diagnosa Penyakit Covid-19

Guidio Leonarde Ginting, Natalia Silalahi*

Fakultas Ilmu Komputer dan Teknologi Informasi, Universitas Budi Darma, Medan, Indonesia

Email: ¹guidio.leonard626@gmail.com, ^{2,*}natalia.novena.silalahi@gmail.com

INFORMASI ARTIKEL

Article History:

Received : Jun 29, 2022

Accepted : Jul 21, 2022

Published : Jul 31, 2022

KORESPONDENSI

Email: natalia.novena.silalahi@gmail.com

A B S T R A K

Covid-19 merupakan sebuah virus yang berukuran sangat kecil sehingga tidak dapat dilihat dengan kasat mata, di perlukan adanya bantuan alat pembesar seperti mikroskop dan virus ini sangat berbahaya yang mengakibatkan gangguan pernafasan pada manusia bahkan dapat mematikan manusia yang telah terpapar virus covid-19 atau virus corona ini. Dimana Pada umumnya masyarakat maupun para tim medis dalam mendiagnosa seseorang yang telah terpapar Covid -19 berdasarkan sebuah gejala antara lain: demam, batuk, pilek, sesak nafas, tidak dapat mencium bau dan merasakan pada lidah. Tetapi tidak semua orang yang menderita penyakit seperti itu dapat di golongan kedalam penyakit Covid -19. Sehingga untuk memastikan seseorang sedang Covid-19 maka dilakukan test antigen atau test PCR. Data pemeriksaan terkait Covid-19 dari hasil gejala dengan hasil test antigen atau PCR sudah banyak terkumpul dan belum dimanfaatkan untuk mendiagnosa seseorang yang mengalami gejala Covid-19. Dari permasalahan yang ada para tim medis sangat memerlukan bantuan dalam mendiagnosa penyakit covid-19 pada pasien dengan mudah dan cepat yaitu dengan menggunakan data mining algoritma naïve bayes, yang dimana algoritma naïve bayes dapat menghasilkan data yang benar dan nyata. Hasil dari penelitian ini berdasarkan data mining algoritma naïve bayes dalam diagnosa penyakit covid- 19 yaitu Alex positif terkena covid -19 Sedangkan Imam negatif terkena covid-19.

Kata Kunci: Data Mining; Pasien; Covid-19; Algoritma Naïve Bayes

A B S T R A C T

Covid-19 is a virus that is so small that you cannot see it with the naked eye, you need the help of a magnifying device such as a microscope and this virus is very dangerous which causes respiratory problems in humans and can even kill humans who have been exposed to the Covid-19 virus or other viruses. this crown. Where in general, the community and medical teams diagnose someone who has been exposed to Covid-19 based on a symptom including: fever, cough, runny nose, shortness of breath, unable to smell and feel on the tongue. But not everyone who suffers from such a disease can be classified as Covid -19 disease. So to make sure someone has Covid-19, an antigen test or PCR test is carried out. Examination data related to Covid-19 from the results of symptoms with the results of antigen or PCR tests has been collected and has not been used to diagnose someone who has symptoms of Covid-19. From the problems that exist, the medical team really needs help in diagnosing Covid-19 disease in patients easily and quickly, namely by using the Naïve Bayes algorithm data mining, where the Naïve Bayes algorithm can produce correct and real data. The results of this study are based on the Naïve Bayes data mining algorithm in the diagnosis of Covid-19, namely Alex is positive for Covid-19 while Imam is negative for Covid-19.

Keywords: Data Mining; Patients; Covid-19; Naïve Bayes Algorithm

1. PENDAHULUAN

Timbulnya virus covid-19 atau biasanya disebut dengan virus corona berawal Pada tahun 2019 di kota wuhan tepatnya pada negara China, setelah itu meluas di seluruh dunia bahkan di negara Indonesia. Namun pada tahun 2020 indonesia baru saja dilanda oleh virus covid-19 ini. Covid-19 merupakan sebuah virus yang berukuran sangat kecil sehingga tidak dapat dilihat dengan kasat mata sehingga di perlukan adanya bantuan alat pembesar seperti mikroskop dan sangat berbahaya yang mengakibatkan gangguan pernafasan setiap manusia bahkan dapat mematikan manusia yang telah terpapar virus covid-19 atau virus corona ini[1]. Masyarakat pada umum kesulitan dalam membedakan yang terpapar dengan virus Covid-19 dan yang tidak terpapar. Karena gejala yang terpapar dari virus corona atau Covid-19 hampir

sama seperti penyakit pada umumnya, dimana gejalanya yaitu seperti sesak napas, demam, batuk tetapi tidak influenza.

Pada umumnya masyarakat maupun para tim medis dalam mendiagnosa seseorang yang telah terpapar Covid-19 berdasarkan sebuah gejala antara lain: demam, batuk, pilek, sesak napas, Indra Penciuman Tidak Berfungsi dan Indra Perasa Tidak Berfungsi[2]. Namun tidak semua orang yang menderita penyakit seperti itu bisa di golongan kedalam penyakit Covid-19 saja tetapi bisa juga di golongan kedalam penyakit ringan pada umumnya yang mudah di sembuhkan. Sehingga untuk memastikan seseorang sedang Covid-19 maka dilakukan test antigen atau test PCR. Data pemeriksaan terkait Covid-19 dari hasil gejala dengan hasil test antigen atau PCR sudah banyak terkumpul dan belum dimanfaatkan untuk mendiagnosa seseorang yang mengalami gejala Covid-19.

Data mining merupakan suatu kegiatan yang berkaitan dengan proses pengumpulan data informasi dari data yang berkapasitas besar[3]. Ada pula algoritma atau metode yang terdapat dalam data mining seperti algoritma C4.5, algoritma K-Means, algoritma Priori, algoritma hash – based, algoritma K-Medoids, algoritma Naive Bayes, dan lainnya. Dalam penelitian ini penulis menggunakan algoritma Naive Bayes merupakan suatu teknologi dalam metode atau pun algoritma yang penjelasannya sangat mempermudah dalam penerapan bahkan cepat dalam pemrosesan[4]. Metode Naive Bayes bertujuan untuk membantu tim medis dalam menangani pasien dengan cepat dan mudah.

Penulis menjadikan penelitian terkait sebagai titik acuan dalam metode naïve bayes agar mendapatkan hasil yang benar. Dari penelitian yang dilakukan oleh Ratih Yulia Hayuningtyas pada tahun 2019 Dengan judul penerapan algoritma naïve bayes untuk rekomendasi pakaian wanita, dari penelitian ini penulis menyimpulkan data mining dengan menggunakan algoritma naïve bayes mempermudah dan dapat memecahkan masalah dalam pengambilan keputusan rekomendasi pakaian wanita [5]. Penelitian yang lain yang dilakukan oleh Mochamad Farid Rifai, Hendra Jatnika, Bowval Valentino pada tahun 2019. Yang berjudul penerapan algoritma naïve bayes pada sistem prediksi prestasi tingkat kelulusan peserta sertifikasi Microsoft office specialist (MOS), menyimpulkan hasil dari penelitian ini 0,001042 sebesar tingkat akurasi 99% dengan nilai probabilitas untuk memprediksi tingkat kelulusan peserta sertifikat MOS [6].

2. METODOLOGI PENELITIAN

2.1 Data Mining

Data mining ialah sekumpulan data informasi yang telah dilakukan penggalian atau dikumpulkan data dari berkapasitas besar hingga berkapasitas kecil[7]. Data mining juga bertujuan sebagai menentukan keputusan dari suatu bisnis yang penting. Dari penelitian ini menggunakan data mining sebab para tim medis kesulitan atau kewalahan dalam menangani pasien sehingga tim medis menggunakan data mining untuk mempermudah menanalisa dari data – data pasien bagi para tim medis[8].

2.2 Pasien

Pasien merupakan seseorang yang menderita penyakit yang dimana penyakit tersebut harus adanya pengawasan dari tim medis seperti dokter dan suster[9]. Selain itu juga pasien ialah seseorang yang sedang berobat dan melakukan rawat inap di sebuah rumah sakit.

2.3 Covid-19

Covid-19 adalah virus yang bisa merusak organ tubuh dan dapat mengakibatkan fatal sehingga dapat membuat seseorang meninggal dunia[10]. Adapun cara dalam memutuskan rantai virus covid-19 atau sering disebut corona yaitu seperti rajin cuci tangan, tetap menggunakan masker, tetap menjaga jarak dan menjauhi yang namanya keramaian[1].

2.4 Algoritma Naïve bayes

Thomas Bayes adalah seorang ilmuwan yang telah menemukan naïve bayes ialah suatu pengklasifikasian dengan menggunakan metode statistik dan probabilitas. Naive bayes merupakan sebuah metode atau algoritma yang terdapat dari berbagai banyaknya algoritma data mining yang didalam teknologi klasifikasi yang sangat mudah dan cepat dalam pemrosesan data agar dapat di implementasikan[11]. Adapun rumus dari algoritma naïve bayes antara lain yaitu[12]:

$$P(H|X) = \frac{P(X|H)(PH)}{P(X)} \quad (1)$$

Dimana:

X = suatu data sampel yang mempunyai label (kelas)

H = sebuah hipotensi x pada data label (kelas)

P(H) = suatu peluang pada hipotesis H

P(X) = sebuah peluang data sampel yang mempunyai label (kelas)

P(X|H) = Sebuah peluang pada sampel data X dari sebuah data hipotesis data label

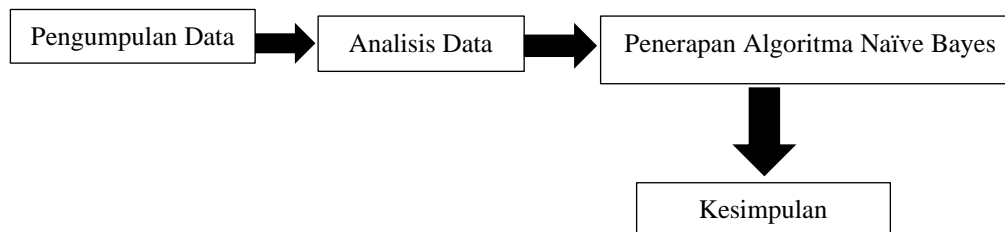
Cara kerja dari metode algoritma naïve bayes sebagai berikut:

1. terlebih dahulu siapkan sebuah data yang dalam bentuk kasus lama (decision system)
2. langkah kedua, siapkan juga kasus baru atau data baru
3. Langkah ketiga, hitunglah sebuah peluang hipotesa P(X|H) pada setiap kriteria

4. Langkah keempat, hitung juga $P(X|H)$ $P(H)$ pada setiap label atau kelas
5. Langkah kelima, hitunglah $(P(X|H)(P(H))/(P(X))$
6. Langkah terakhir, hasil (bandingkanlah nilai antar setiap kelas atau label, nilai tertinggi yang akan menjadi hasilnya).

2.5 Tahapan Penelitian

Didalam penelitian terdapat sebuah metodologi penelitian, metodologi penelitian merupakan suatu tahapan penyusunan penelitian yang dilakukan oleh penulis. Metodologi penelitian dalam sebuah penelitian yang bertujuan dapat memudahkan para penulis lainnya dan penulis itu sendiri dalam Menyusun sebuah penelitian. Dipenelitian ini Adapun tahapan penyusunan yang diambil dari metodologi penelitian antara lain:



Gambar 1. Kerangka Penelitian

Adapun beberapa penjelasan singkat pada tahap penelitian tersebut sebagai berikut.

1. Pengumpulan Data

Pengumpulan data adalah suatu cara atau Langkah yang di gunakan penulis dalam mengumpulkan sebuah informasi yang ada pada area lapangan. Penelitian ini memiliki beberapa cara dalam mengumpulkan data ataupun informasi sebagai berikut:

a. Observasi

Dari observasi atau pengamatan pada penelitian ini, penulis melakukan pengamatan secara langsung agar mendapatkan sebuah informasi dan data yang diinginkan penulis.

b. Wawancara

Pada penelitian ini penulis langsung menanyakan atau mewawancarai pihak yang bersangkutan seperti pada pihak rumah sakit. Wawancara bertujuan untuk memperoleh data dan sebuah informasi.

2. Studi literatur

Studi literatur sangat diperlukan untuk penelitian supaya penulis dapat menerima dan dimengerti dari suatu konsep data mining algoritma yang lebih rinci dan memahami cara perhitungan algoritma naïve bayes. Studi literatur ini juga penulis menjadikan sebagai referensi pada penelitian ini.

3. Analisis Data

Penulis melakukan analisis terhadap tahapan atau cara dari algoritma naïve bayes dalam melakukan sebuah proses mendiagnosa penyakit covid – 19.

4. Penerapan algoritma naïve bayes

Pada penelitian ini penulis menggunakan algoritma naïve bayes yang bertujuan dapat membantu dan mempermudah tim rumah sakit dalam mendiagnosa penyakit covid – 19.

5. Kesimpulan

Tahapan yang terakhir yaitu kesimpulan, dimana kesimpulan terdapat hasil dari permasalahan yang sudah di jabarkan diatas.

3. HASIL DAN PEMBAHASAN

3.1 Data Pasien Covid 19

Dalam data mining bukan hanya memerlukan data kriteria saja, namun harus ada juga data alternatif untuk mempermudah diagnosa penyakit covid-19 pada pasien. Pada table 1 dan 2 adalah data pasien dibawah ini yaitu:

Tabel 1. Data Pasien Lama

No	Nama Pasien
1	Gunawan
2	Faisal
3	Dandi
4	Irma
5	Bayu
6	Yuda
7	Putri
8	Salsabila
9	Dwi

No	Nama Pasien
10	Agung
11	Dinda
12	Dodi
13	Ira
14	Joko
15	Susanti
16	Sinta
17	Wulan
18	Panji
19	Fandi
20	Irfan

Tabel 2. Data Pasien Baru

No	Nama Pasien
1	Alex
2	Imam

Untuk mendiagnosa penyakit covid 19 pada pasien dalam data mining harus adanya data yang mendukung agar menghasilkan data yang nyata. Pada tabel 3 Data Baru dan Data Lama Pasien seperti di bawah ini:

Tabel 3 Data Lama Pasien

Nama Pasien	Demam	Batuk	Filek	Sesak Nafas	Indra penciuman tidak berfungsi	Indra Perasa Tidak Berfungsi	Ket
Gunawan	Iya	Tidak	Iya	Iya	Tidak	Iya	Positif
Faisal	Iya	Iya	Iya	Iya	Tidak	iya	Positif
Dandi	Tidak	Iya	Tidak	Iya	Iya	Iya	Positif
Irma	Iya	Tidak	Iya	Tidak	Tidak	Tidak	Negatif
Bayu	Iya	Tidak	Iya	Iya	Tidak	Iya	Positif
Yuda	Iya	Tidak	Iya	Tidak	Tidak	Tidak	Negatif
Putri	Tidak	Iya	Tidak	Tidak	Iya	Tidak	Negatif
Salsabila	Tidak	Iya	Tidak	Tidak	Iya	Tidak	Negatif
Dwi	Iya	Tidak	Iya	Iya	Tidak	Iya	Positif
Agung	Tidak	Iya	Tidak	Iya	Iya	Iya	Positif
Dinda	Iya	Tidak	Iya	Iya	Tidak	Tidak	Negatif
Dodi	Tidak	Iya	Tidak	Iya	Iya	Iya	Positif
Ira	Iya	Tidak	Iya	Tidak	Tidak	Tidak	Negatif
Joko	Iya	Tidak	Iya	Iya	Tidak	Iya	Positif
Susanti	Iya	Tidak	Iya	Tidak	Tidak	Tidak	Negatif
Sinta	Tidak	Iya	Tidak	Iya	Iya	Iya	Positif
Wulan	Tidak	Iya	Tidak	Iya	Iya	Iya	Positif
Panji	Tidak	Iya	Tidak	Tidak	Iya	Tidak	Negatif
Fandi	Iya	Tidak	Iya	Tidak	Tidak	Tidak	Negatif
Irfan	Iya	Tidak	Iya	Tidak	Tidak	Tidak	Negatif

Keterangan:

Demam : suatu kenaikan suhu tubuh yang panasnya diatas 38 derajat Celsius.
 Batuk : suatu respons dari tubuh yang berguna untuk mengeluarkan zat.
 Filek : suatu hidung yang mengeluarkan lender atau ingus.
 Sesak Nafas : seorang yang sedang mengalami gejala kesulitan bernafas.
 Indra Penciuman Tidak Berfungsi : penyakit yang tidak bisa mencium bau secara stabil.
 Indra Perasa Tidak Berfungsi : penyakit yang tidak bisa merasakan dengan stabil

Tabel 4. Data Baru Pasien

Nama Pasien	Demam	Batuk	Filek	Sesak Nafas	Indra Penciuman Tidak Berfungsi	Indra Perasa Tidak Berfungsi	Ket
Alex	Iya	Tidak	Iya	Iya	Iya	Iya	?
Imam	Iya	Tidak	Iya	Tidak	Tidak	Tidak	?

Penyelesaian:

- Menghitung $P(X|H)$ untuk setiap kriteria
 - $P(\text{Demam} = \text{"Iya"} \mid \text{keterangan} = \text{"positif"})$
 $P(5/10 = 0,5)$
 - $P(\text{Demam} = \text{"Iya"} \mid \text{keterangan} = \text{"negatif"})$
 $P(7/10 = 0,7)$

3. P (Batuk = "Tidak" | keterangan = "positif")
 $P(4/10 = 0,4)$
4. P (Batuk = "Tidak" | keterangan = "negatif")
 $P(7/10 = 0,7)$
5. P (Filek = "Iya" | keterangan = "positif")
 $P(5/10 = 0,5)$
6. P (Filek = "Iya" | keterangan = "negatif")
 $P(7/10 = 0,7)$
7. P (Sesak Nafas = "Iya" | keterangan = "positif")
 $P(10/10 = 1)$
8. P (Sesak Nafas = "Iya" | keterangan = "negatif")
 $P(1/10 = 0,1)$
9. P (Indra Penciuman Tidak Berfungsi = "Iya" | keterangan = "positif")
 $P(5/10 = 0,5)$
10. P (Indra Penciuman Tidak Berfungsi = "Iya" | keterangan = "negatif")
 $P(3/10 = 0,3)$
11. P (Indra Perasa Tidak Berfungsi = "Iya" | keterangan = "positif")
 $P(10/10 = 1)$
12. P (Indra Perasa Tidak Berfungsi = "Iya" | keterangan = "negatif")
 $P(0/10 = 0)$

b. Menghitung $P(X|H) P(H)$ untuk label

Label = "positif"

$$= 0,5 \times 0,4 \times 0,5 \times 1 \times 0,5 \times 1 = 0,05$$

Label = "negatif"

$$= 0,7 \times 0,7 \times 0,7 \times 0,1 \times 0,3 \times 0 = 0$$

c. Hitung $\frac{P(X|H) P(H)}{P(X)}$ berdasarkan label

Label = $p(x) = (10/20)$ – "positif"

$$= 0,05 / (10/20)$$

$$= 0,05 / 0,5$$

$$= 0,1$$

Label = $p(x) = (10/20)$ – "negatif"

$$= 0 / (10/20)$$

$$= 0 / 0,5$$

$$= 0$$

Berdasarkan nilai label "positif" lebih tinggi dari pada label "negatif", maka dapat disimpulkan bahwa Alex positif terkena covid -19.

a. Menghitung $p(X|H)$ untuk setiap kriteria

1. P (Demam = "Iya" | Keterangan = "Positif")
 $P(5/10 = 0,5)$
2. P (Demam = "Iya" | keterangan = "negatif")
 $P(7/10 = 0,7)$
3. P (Batuk = "Tidak" |Keterangan = "positif")
 $P(4/10 = 0,4)$
4. P (Batuk = "Tidak" |Keterangan = "negatif")
 $P(7/10 = 0,7)$
5. P (Filek = "Iya" |Keterangan = "positif")
 $P(5/10 = 0,5)$
6. P (Filek = "Iya" |Keterangan = "negatif")
 $P(7/10 = 0,7)$
7. P (Sesak Nafas = "Tidak" |Keterangan = "positif")
 $P(10/10 = 1)$
8. P (Sesak Nafas = "Tidak" |Keterangan = "negatif")
 $P(1/10 = 0,1)$
9. P (Indra Penciuman Tidak Berfungsi = "Tidak" |Keterangan = "positif")
 $P(5/10 = 0,5)$
10. P (Indra Penciuman Tidak Berfungsi = "Tidak" |Keterangan = "negatif")
 $P(3/10 = 0,3)$
11. P (Indra Perasa Tidak Berfungsi = "Tidak" |Keterangan = "positif")
 $P(10/10 = 1)$
12. P (Indra Perasa Tidak Berfungsi = "Tidak" |Keterangan = "negatif")
 $P(0/10 = 0)$

b. Mengitung Menghitung $P(X|H) P(H)$ untuk label

Label = "positif"

$$= 0,5 \times 0,4 \times 0,5 \times 0 \times 0,5 \times 0 = 0$$

Label = “negatif”

$$= 0,7 \times 0,7 \times 0,7 \times 0,9 \times 0,7 \times 1 = 0,21609$$

c. Hitung $\frac{P(X|H)(PH)}{P(X)}$ berdasarkan label

Label = $p(x) = (10/20)$ – “positif”

$$= 0 / (10/20)$$

$$= 0/0,5$$

$$= 0$$

Label = $p(x) = (10/20)$ – “negatif”

$$= 0,21609 / (10/20)$$

$$= 0,21609 / 0,5$$

$$= 0,43218$$

Berdasarkan nilai label “negatif” lebih tinggi dari pada label “positif”, maka dapat disimpulkan bahwa Imam negatif terkena covid-19.

4. KESIMPULAN

Dari permasalahan yang telah dijelaskan sehingga penulis dapat menyimpulkan dari hasil penelitian ini yaitu antara lain, algoritma naïve bayes pada data mining sangat berguna dan membantu para tim medis di rumah sakit untuk mendiagnosa pasien yang terkena atau terpapar dari virus covid -19. Dimana hasil dari algoritma naïve bayes adalah hasil yang mutlak dan tidak ada siapapun yang bisa mengubahnya. Hasil dari penelitian ini berdasarkan data mining algoritma naïve bayes dalam diagnosa penyakit covid- 19 yaitu Alex positif terkena covid -19 Sedangkan Imam negatif terkena covid-19.

REFERENCES

- [1] Y. Putri, “IMBAUAN MEMAKAI MASKER PADA MASA PANDEMI COVID-19 DI PERUMAHAN GARUDA TELUKNAGA TANGERANG (Perspektif Psikologi Komunikasi),” *SIGI*, vol. 1, no. 1, pp. 49–57, 2019.
- [2] K. T. Dewi, A. F. Boy, and S. Murniyanti, “Sistem Pendukung Keputusan Menentukan Rumah Sakit Terbaik Dalam Pelayanan Pasien COVID-19 Menggunakan Metode ARAS,” *J. Cyber Tech*, vol. 2, no. 8, 2019.
- [3] N. Sirait, “IMPLEMENTASI K-MEANS CLUSTERING PADA PENGELOMPOKAN MUTU BIJI SAWIT (Studi Kasus : PT . Multimas Nabati Asahan),” *J. Pelita Inform.*, vol. 6, no. 2, pp. 170–174, 2017.
- [4] A. A. Rahman, Y. I. Kurniawan, P. S. Informatika, and U. M. Surakarta, “APLIKASI KLASIFIKASI PENERIMA KARTU INDONESIA SEHAT MENGGUNAKAN,” 2016.
- [5] R. Y. Hayuningtyas, “Penerapan Algoritma Naïve Bayes untuk Rekomendasi Pakaian Wanita,” *J. Inform.*, vol. 6, no. 1, pp. 18–22, 2019, doi: 10.31311/ji.v6i1.4685.
- [6] M. F. Rifai, H. Jatnika, and B. Valentino, “Penerapan Algoritma Naïve Bayes Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist (MOS),” *Petir*, vol. 12, no. 2, pp. 131–144, 2019, doi: 10.33322/petir.v12i2.471.
- [7] C. Romero and S. Ventura, “Data mining in education,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 3, no. 1, pp. 12–27, 2013.
- [8] M. Huda and M. Kom, *Algoritma Data Mining: Analisis Data Dengan Komputer*. bisakimia, 2019.
- [9] D. A. Sitepu, “Tanggung Jawab Keperdataan Dokter Muda Dalam Penanganan Pelayanan Kesehatan Terhadap Pasien Di Rumah Sakit.” 2019.
- [10] P. Covid, “di Indonesia,” *Retrieved Februari*, vol. 25, p. 2021, 19AD.
- [11] D. Laia, E. Buulolo, and M. J. F. Sirait, “Implementasi Data Mining Untuk Memprediksi Pemesanan Driver Go-Jek Online Dengan Menggunakan Metode Naive Bayes (Studi Kasus: Pt. Go-Jek Indonesia),” *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 2, no. 1, pp. 434–439, 2018, doi: 10.30865/komik.v2i1.972.
- [12] H. Zhang, “The optimality of naive Bayes,” *Aa*, vol. 1, no. 2, p. 3, 2004.