

**ANALISIS SENTIMEN TWITTER VAKSIN COVID-19
MENGUNAKAN KOMBINASI *NAÏVE BAYES CLASSIFIER*
DAN *INFORMATION GAIN***

SKRIPSI

NIEKA JULYANA

20180040116



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK KOMPUTER DAN DESAIN
SUKABUMI
JULI 2022**

**ANALISIS SENTIMEN TWITTER VAKSIN COVID-19
MENGUNAKAN KOMBINASI *NAÏVE BAYES CLASSIFIER*
DAN *INFORMATION GAIN***

SKRIPSI

NIEKA JULYANA

20180040116



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK KOMPUTER DAN DESAIN
SUKABUMI
JULI 2022**

PERNYATAAN PENULIS

JUDUL : ANALISIS SENTIMEN TWITTER VAKSIN COVID-19
MENGGUNAKAN KOMBINASI NAÏVE BAYES
CLASSIFIER DAN INFORMATION GAIN

NAMA : NIEKA JULYANA

NIM : 20180040116

“Saya menyatakan bertanggung jawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Komputer saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut”.

Sukabumi, 14 Juli 2022



PERSETUJUAN SKRIPSI

JUDUL : ANALISIS SENTIMEN TWITTER VAKSIN COVID-19
MENGGUNAKAN KOMBINASI NAÏVE BAYES
CLASSIFIER DAN INFORMATION GAIN

NAMA : NIEKA JULYANA

NIM : 20180040116

Skripsi ini telah diperiksa dan disetujui

Sukabumi, 14 Juli 2022

Ketua Progran Studi
Teknik Informatika,

Anggun Fregina, M.Kom

NIDN. 0407029301



Library Innovation Unit
LIU

Pembimbing,

Alun Sujjada, S.Kom, M.T

NIDN. 0718108001

PENGESAHAN SKRIPSI

JUDUL : ANALISIS SENTIMEN TWITTER VAKSIN COVID-19
MENGGUNAKAN KOMBINASI NAÏVE BAYES
CLASSIFIER DAN INFORMATION GAIN

NAMA : NIEKA JULYANA

NIM : 20180040116

Skripsi ini telah diujikan dan dipertahankan didepan Dewan Penguji pada Sidang Skripsi tanggal 11 Juli 2022 Menurut pandangan kami. Skripsi ini memadai dari segi kualitas untuk tujuan penganugrahan gelar Sarjana Komputer (S.Kom).

Sukabumi, 14 Juli 2022

Pembimbing I

Alun Sujjada, S.Kom, M.T
NIDN. 0718108001

Ketua Penguji

Kamdan, S.T, M.Kom
NIDN. 0401107401

Pembimbing II

Gina Purnama Insany, M.Kom
NIDN. 0717077908

Ketua Program Studi

Anggun Fergina, M.Kom
NIDN. 0407029301

Dekan Fakultas Teknik Komputer Dan Desain

Prof. Dr. Ir. H. Koesmawan, M.Sc. MBA, DBA
NIDN. 0014075205

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIK

Sebagai civitas akademik Universitas Nusa Putra, saya yang bertanda tangan di bawah ini:

Nama : Nieka Julyana
NIM : 20180040116
Program Studi : Teknik Informatika
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Nusa Putra **Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty- Free Right)** atas karya ilmiah kami yang berjudul :

ANALISIS SENTIMEN TWITTER VAKSIN COVID-19 MENGGUNAKAN KOMBINASI NAÏVE BAYES CLASSIFIER DAN INFORMATION GAIN

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Nusa Putra berhak menyimpan, mengalihmedia / format-kan, mengelola dalam bentuk pangkalan data (database), merawat dan memublikasikan tugas akhir saya selama tetap mencantumkan nama kami sebagai penulis / pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Sukabumi
Pada Tanggal : 14 Juli 2022

Yang menyatakan,

NIEKA JULYANA
NIM. 20180040116

HALAMAN PERUNTUKAN

Alhamdulillahirabbilalamin ...

Ucap syukur kehadiran Illahi Rabbi, karena berkat rahmat dan Hidayah-Nya saya dapat menyelesaikan Skripsi ini dengan segala kekurangannya. Skripsi ini diperuntukan kepada Mamah, Bapak, Kakak-kakakku tersayang yang tiada hentinya memberikan kasih sayang, dan dukungan do'a serta kepada diriku sendiri yang telah mau berjuang sampai sejauh ini.



ABSTRAK

Covid-19 telah membawa dampak perubahan dalam aspek kehidupan dan sosial masyarakat. Pemerintah melakukan strategi untuk melawan kasus tersebut dengan melakukan vaksinasi kepada masyarakat. Kebijakan ini mewajibkan masyarakat untuk mengikuti kegiatan vaksinasi tersebut dari mulai vaksin ke-1, vaksin Ke-2, sampai vaksin yang ke-3. Namun dari kegiatan vaksinasi tersebut kemudian menjadi kontroversi dikarenakan banyaknya isu yang beredar mengenai efek setelah melakukan vaksin dengan hal itulah yang menjadikan masyarakat enggan melakukan vaksinasi dan memberikan berbagai respon atau opini di masyarakat. Maka dari itu penelitian ini bertujuan untuk melakukan analisis sentimen mengenai vaksin Covid-19 dengan menggunakan algoritma *Naïve Bayes* dan seleksi fitur *Information Gain*. Hasil dari penerapan algoritma *Naïve Bayes* dengan kombinasi seleksi fitur *Information Gain* menunjukkan bahwa tingkat akurasi yang optimal yakni nilai akurasi mencapai 90.53%, presisi 93.57%, serta *recall* 98.31% dari data uji yang diambil sebanyak 500 data.

Kata kunci : Vaksin Covid-19, Analisis Sentimen, *Naïve Bayes*, *Information Gain*



ABSTRACT

Covid-19 has brought changes in aspects of life and social society. The government carried out a strategy to fight this case by vaccinating the community. This policy requires the public to participate in the vaccination activities starting from the 1st vaccine, 2nd vaccine, until the 3rd vaccine. However, the vaccination activity became controversial due to the many issues circulating regarding the effects of having the vaccine, which made people reluctant to vaccinate and gave various responses or opinions in the community. Therefore, this study aims to analyze sentiment regarding the Covid-19 vaccine by using the Naïve Bayes algorithm and selecting the Information Gain feature. The results of the application of the Naïve Bayes algorithm with a combination of Information Gain feature selection show that the optimal level of accuracy is the accuracy value reaches 90.53%, precision is 93.57%, and recall is 98.31% from 500 test data taken.

Keywords : Covid-19 Vaccine, Sentiment Analysis, Nave Bayes, Information Gain



KATA PENGANTAR

Puji syukur panjatkan kehadirat Allah SWT, berkat Rahmat, Hidayah dan Karunia-Nya sehingga peneliti dapat menyelesaikan skripsi dengan judul “Analisis Sentimen *Twitter* Vaksin Covid-19 Menggunakan Kombinasi *Naïve Bayes Classifier* dan *Information Gain*” dengan baik. Shalawat serta salam terlimpah curah kepada Nabi Muhammad SAW, beserta sahabat dan keluarganya.

Adapun maksud dan tujuan penulisan Laporan Skripsi ini adalah sebagai gambaran terhadap apa yang peneliti kerjakan. Selain itu juga laporan ini sebagai salah satu syarat memperoleh Gelar Sarjana pada Program Studi Teknik Informatika di Universitas Nusa Putra.

Penulis menyadari dalam penyusunan skripsi ini tidak lepas dari dukungan dan partisipasi dari berbagai pihak, sehingga penulis dapat menyelesaikannya dengan baik dan lancar. Oleh karena itu, pada kesempatan ini kami ingin mengucapkan banyak terima kasih yang sebesar-besarnya, kepada :

1. Bapak Dr. Kurniawan S.T, M.Si, M.M., selaku Rektor Universitas Nusa Putra Sukabumi.
2. Bapak Anggy Pradiftha Junfithrana, S.Pd, M.T, selaku Wakil Rektor I Bidang Akademik Universitas Nusa Putra.
3. Ibu Anggun Fergina, M.Kom, selaku Ketua Program Studi Teknik Informatika Universitas Nusa Putra Sukabumi.
4. Bapak Alun sujjada, S.Kom, M.T, selaku Dosen Pembimbing I yang telah meluangkan waktu, tenaga dan pikiran untuk mengarahkan saya dalam menyelesaikan skripsi ini.
5. Ibu Gina Purnama Insany, M.Kom selaku pembimbing II yang telah memberikan *support* juga telah meluangkan waktu untuk mengarahkan saya dalam penulisan penyusunan skripsi ini.
6. Dosen Penguji Skripsi, terimakasih atas saran dan masukan yang diberikan.
7. Seluruh Dosen Jurusan Teknik Informatika Universitas Nusa Putra yang tidak dapat disebutkan satu persatu yang telah berjasa memberikan ilmu pengetahuan.

8. Kepada orang tua saya, Ibu Iis Aisyah Dan Bapak Dadang Hodori yang tiada henti memberikan do'a, dukungan, mendidik, serta memberikan cinta yang tulus kepada saya semenjak kecil hingga saat ini baik materil maupun non materil.
9. Teman-teman seperjuangan yang selalu memberikan semangat dan selalu menemani dari awal perkuliahan sampai sekarang.

Penulis menyadari sepenuhnya dalam penyusunan skripsi ini masih banyak terdapat kekurangan mengingat keterbatasan dalam pengetahuan dan kemampuan, walaupun penulis telah berusaha dengan sebaik-baiknya. Oleh karena itu demi perkembangan penelitian selanjutnya penulis sangat mengharapkan kritik dan saran yang sifatnya membangun. Akhir kata, semoga skripsi ini dapat bermanfaat khususnya bagi penulis dan juga bagi para pembaca.

Sukabumi, Juni 2022

Penulis



DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
PERNYATAAN PENULIS	iii
PERSETUJUAN SKRIPSI.....	iv
PENGESAHAN SKRIPSI.....	v
HALAMAN PERNYATAAN PUBLIKASI	vi
HALAMAN PERUNTUKAN	vii
ABSTRAK	viii
ABSTRACT	ix
KATA PENGANTAR.....	x
DAFTAR ISI.....	xii
DAFTAR GAMBAR.....	xiv
DAFTAR TABEL.....	xv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah.....	4
1.4. Tujuan Penelitian.....	4
1.5. Manfaat Penelitian.....	4
1.6. Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA.....	7
2.1. Penelitian Terkait	7
2.2. Data Mining.....	9
2.3. <i>Text Mining</i>	10
2.4. Analisis Sentimen.....	11
2.5. <i>Twitter</i>	12
2.6. <i>R Programming</i>	12
2.7. <i>Regular Expression</i>	13
2.8. <i>Lexical BASED</i>	15
2.9. <i>PHP</i>	15
2.10. <i>MySQL</i>	16
2.11. <i>Preprocessing</i>	17

2.12. <i>Supervised Learning</i>	18
2.13. <i>Information Gain</i>	18
2.14. Naïve Bayes Classifier	19
2.15. Evaluasi	21
2.16. <i>Flowchart</i>	22
2.17. Kerangka Berpikir	24
BAB III METODOLOGI PENELITIAN	25
3.1. Metodologi Penelitian	25
3.2. Pengumpulan Data	26
3.3. Pengambilan Data	26
3.4. Preprocessing Data	28
3.5. Pelabelan Data.....	35
3.6. Seleksi <i>Fitur Information Gain</i>	35
3.7. Pembagian Data.....	36
3.8. Klasifikasi Naïve Bayes	37
3.9. <i>Evaluasi</i>	39
BAB IV HASIL DAN PEMBAHASAN	40
4.1. Pengambilan Data	40
4.2. <i>Preprocessing Data</i>	41
4.3. Pelabelan Data.....	43
4.4. Seleksi <i>Fitur Information Gain</i>	44
4.5. Klasifikasi <i>Naïve Bayes</i>	45
4.6. Desain Dan Implementasi GUI	46
4.7. Evaluasi	50
4.8. Visualisasi	58
BAB V PENUTUP	59
5.1. Kesimpulan.....	59
5.2. Saran.....	60
DAFTAR PUSTAKA	61
LAMPIRAN	64

DAFTAR GAMBAR

Gambar 2. 1 Tahapan Data Mining.....	10
Gambar 2. 2 Tahapan Text Mining.....	11
Gambar 2. 3 Contoh Sintaks R.....	13
Gambar 3. 1. Tahap Penelitian.....	25
Gambar 3. 2 Pengambilan Data	27
Gambar 3. 3 Sintaks Proses Pengambilan data	27
Gambar 3. 4 Preprocessing Data	29
Gambar 3. 5 Flowchart Cleanning	30
Gambar 3. 6 Flowchart Stopword	31
Gambar 3. 7 Flowchat Tokenizing.....	33
Gambar 3. 8 Flowchat Case Folding.....	34
Gambar 3. 9 Flowchart Penghitungan Klasifikasi Naive Bayes	37
Gambar 4. 1 Tampilan Twitter Developer	40
Gambar 4. 2 Proses Crawling Data	40
Gambar 4. 3 Hasil Crawling.....	41
Gambar 4. 4 Proses Cleanning	41
Gambar 4. 5 Proses Stopword.....	42
Gambar 4. 6 Proses Stemming	42
Gambar 4. 7 Proses Tokenizing	42
Gambar 4. 8 Proses Case Folding	43
Gambar 4. 9 Hasil Word Cloud.....	43
Gambar 4. 10 Visualisasi Label	44
Gambar 4. 11 Tampilan Halaman Utama	46
Gambar 4. 12 Tampilan Halaman data Hasil Preprocessing.....	47
Gambar 4. 13 Tampilan Halaman Label Data Training.....	47
Gambar 4. 14 Tampilan Halaman Bobot Information Gain	48
Gambar 4. 15 Tampilan Pengujian Hasil Klasifikasi.....	48
Gambar 4. 16 Tampilan Halaman Visualisasi.....	49
Gambar 4. 17 Hasil Perhitungan Confusion Matrix 125 Data	50
Gambar 4. 18 Hasil Evaluasi 250 Data	51
Gambar 4. 19 Hasil Perhitungan Confusion Matrix 250 Data	51
Gambar 4. 20 Hasil Evaluasi 250 Data	52
Gambar 4. 21 Hasil Perhitungan Confusion Matrix 300 Data	52
Gambar 4. 22 Hasil Evaluasi 300 Data	53
Gambar 4. 23 Hasil Perhitungan Confusion Matrix 450 Data	53
Gambar 4. 24 Hasil Evaluasi 450 Data	54
Gambar 4. 25 Hasil Perhitungan Confusion Matrix 500 Data	54
Gambar 4. 26 Hasil Evaluasi 500 Data	55
Gambar 4. 27 Hasil Perhitungan CM Keseluruhan Data	56
Gambar 4. 28 Hasil Evaluasi Keseluruhan Data	56
Gambar 4. 29 Visualisasi Pengujian	58

DAFTAR TABEL

Tabel 2. 1 Regular Expression Kelas Karakter	13
Tabel 2. 2 Sintaks Dasar PHP	15
Tabel 2. 3 Contoh Sintaks Dasar MySQL.....	16
Tabel 2. 4 Jenis Flowchart	22
Tabel 3. 1 Regular Expresion Pemrograman R.....	28
Tabel 3. 2 Contoh Tahap Cleanning	31
Tabel 3. 3 Contoh Tahap Stopword	32
Tabel 3. 4 Contoh Tahap Stemming	32
Tabel 3. 5 Contoh Tahap Tokenizing.....	33
Tabel 3. 6 Contoh Tahap <i>Case Folding</i>	34
Tabel 3. 7 Contoh Pelabelan Data.....	35
Tabel 3. 8 Contoh Koleksi Data.....	35
Tabel 3. 9 Contoh Data Training.....	38
Tabel 3. 10 Contoh Data Testing	38
Tabel 4. 1 Hasil Perhitungan Information Gain	45
Tabel 4. 2 Hasil Proses Klasifikasi Naive Bayes	45
Tabel 4. 3 Hasil Perhitungan Data Uji Vaksin Sinovac	55



BAB I

PENDAHULUAN

1.1. Latar Belakang

Corona virus merupakan kumpulan virus yang menyerang sistem pernapasan manusia dari mulai infeksi pernapasan ringan (flu) sampai infeksi pernafasan berat (pneumonia) [1]. Penyebaran virus tersebut dimulai pada akhir tahun 2019, hingga saat ini telah menyebar ke berbagai Negara. Adapun kedatangan virus tersebut ke Indonesia dimulai pada tanggal 14 februari 2020. Data [JHU CSSE COVID-19 Data](#) dan [Our World in Data](#) memperlihatkan bahwa jumlah kasus dari awal kemunculannya hingga saat ini di Indonesia mempunyai total jumlah kasus 4,33 juta dan jumlah kematian yang mencapai 144 rb jiwa. Kepesatan pertumbuhan kasus tersebut telah menimbulkan kepanikan, kecemasan dan ketakutan pada kalangan masyarakat [2]. Efek dari wabah tersebut telah membawa dampak perubahan dalam aspek kehidupan dan sosial masyarakat. Pemerintah pun telah mencoba berbagai cara untuk mengatasi kasus tersebut agar tidak selalu ada peningkatan, salah satu caranya yakni melakukan vaksinasi terhadap masyarakat. Vaksinasi dilakukan agar masyarakat terlindungi dari penyebaran virus corona dan dapat menurunkan tingkat penyebaran virus tersebut.

Ada berbagai varian vaksin yakni *Asrazeneca*, *Sinovac*, *Moderna*, dan lain-lain. Kegiatan vaksinasi untuk masyarakat di Indonesia telah menimbulkan kontroversi hingga mengundang banyak kalangan memberikan opini dan pendapat. Adapun opini dan pendapat tersebut diutarakan dengan berbagai cara, salah satunya di media sosial twitter.

Twitter merupakan jejaring sosial atau media sosial yang banyak diminati oleh masyarakat dikarenakan kemudahan dalam pengaksesannya dan setiap pengguna twitter bebas melakukan posting berbagai *tweet*. *Tweet* merupakan text status yang digunakan untuk memberikan informasi di twitter dari mulai percakapan, berbagi informasi, dan pelaporan berita informasi [3].

Oleh karena itu yang menjadikan alasan mengapa twitter menjadi ladang untuk memberikan opini dan pendapat masyarakat dari berbagai hal, salah satunya mengenai vaksinasi *Covid-19*. Banyak sekali masyarakat yang berpendapat atau memberikan tanggapan mengenai vaksinasi yang dilakukan pemerintah, dari mulai tanggapan negatif dan positif. Maka dari itu, diperlukan analisis sentimen untuk menganalisis opini/pendapat mengenai vaksin *Covid-19* yang dilakukan oleh pemerintah. Alasan penulis mengapa memilih analisis sentimen pada twitter ini guna untuk menganalisa dan mengobservasi perihal isu vaksinasi *Covid-19* yang digencarkan pemerintah kepada masyarakat apakah tanggapan tersebut mayoritas mengarah ke hal positif atau ke hal negatif. Sehingga dapat digunakan masyarakat sebagai bahan pertimbangan untuk mengikuti kegiatan vaksin *Covid-19*.

Analisis sentimen merupakan proses atau teknik yang dipakai untuk melakukan pengklasifikasian emosi dan opini baik positif, netral, dan negatif. Data yang akan dilakukan analisa dikumpulkan dan diambil dari data tweet pengguna twitter. Setelah itu akan dilakukan proses *preprocessing*, pelabelan data secara manual untuk dijadikan data training, lalu pengklasifikasian data menggunakan metode *supervised learning* untuk menentukan apakah postingan tersebut termasuk ke dalam postingan positif, netral atau negatif. Dan untuk metode penghitungan akurasinya yaitu menggunakan metode algoritma Naïve Bayes dengan seleksi fitur Information Gain.

Metode Naïve Bayes ini telah digunakan oleh Evi Dewi Mulyani, dkk yang mana penelitian tersebut menggunakan metode Naïve Bayes sebagai metode klasifikasi untuk analisis sentimen tentang twett mengenai topic pertelevisian di Indonesia dan hasil akurasi pada penelitian tersebut menghasilkan sebesar 91,67%. Dari hasil tersebut telah terbukti bahwa dengan menggunakan algoritma Naïve Bayes ini akan menghasilkan nilai keakurasian yang tinggi [4].

Selanjutnya metode seleksi fitur Information Gain juga telah dilakukan oleh peneliti sebelumnya yakni Sari Widya Sihwi, dkk yang mana metode tersebut digunakan untuk meningkatkan keefisiensian metode Naïve Bayes

dalam mengukur tingkat keakurasian analisis sentiment terhadap ulasan film di twitter dengan tujuan untuk mendapatkan informasi mengenai penilaian/opini masyarakat dari mulai opini negatif, opini positif, dan opini netral. Dari penggunaan metode itu menghasilkan tingkat akurasi pengklasifikasian mencapai 82,19% dan akurasi tersebut merupakan akurasi tertinggi dengan nilai gain threshold 0,006 [5].

Dari hasil kedua penelitian tersebutlah yang menjadikan penulis termotivasi untuk menerapkannya pada analisis sentimen mengenai isu vaksin *covid-19*, karena dengan menggunakan algoritma Naïve Bayes dan seleksi fitur Information Gain akan menghasilkan tingkat akurasi yang optimal. Adapun pembeda atau pembaharuan dalam penelitian ini yaitu dalam pengimplementasian metode *Naïve Bayes* dan *Information Gain* serta untuk menentukan nilai thresholdnya yaitu dengan menggunakan acuan data hasil *labelling*.

Berdasarkan uraian diatas maka dari itu penulis mengambil judul dalam penelitian ini yakni “*Analisis Sentimen Twitter Vaksin Covid-19 Menggunakan Kombinasi Naïve Bayes Classifier dan Information Gain*”.

1.2. Rumusan Masalah

Berdasarkan latar belakang diatas, rumusan masalah pada penelitian ini yaitu:

1. Bagaimana mengkombinasikan antara *Naive Bayes* dan *Information Gain* untuk melakukan analisis sentiment terhadap isu vaksin *Covid-19*?
2. Seberapa optimal hasil akurasi metode *Naïve Bayes Classifier* menggunakan seleksi fitur *Information Gain* dalam pengklasifikasian analisis sentimen Vaksin *Covid-19*?
3. Bagaimana tanggapan masyarakat terhadap kegiatan vaksin yang digencarkan oleh pemerintah

1.3. Batasan Masalah

Adapun batasan masalah dalam melakukan penelitian ini:

1. Data yang diambil untuk dianalisis yaitu dari postingan atau tweet pengguna twitter berbahasa Indonesia
2. Metode yang digunakan untuk pengukuran akurasi adalah *confusion matrix*.

1.4. Tujuan Penelitian

Berdasarkan rumusan masalah diatas maka peneliti merumuskan tujuan penelitian sebagai berikut:

1. Melakukan analisis sentimen mengenai kegiatan vaksin *Covid-19* dengan menggunakan algoritma *Naïve Bayes Classifier* dan seleksi fitur *Information Gain*
2. Mengukur hasil akurasi algoritma *Naïve Bayes Classifier* dan seleksi fitur *Information Gain*

1.5. Manfaat Penelitian

Dari penelitian ini diharapkan dapat bermanfaat bagi pihak-pihak yang memiliki kepentingan mengenai pokok dan pembahasan yang diangkat dari penelitian ini, adapun manfaatnya yaitu:

1. Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan manfaat bagi yang akan melakukan penelitian selanjutnya sebagai masukan perkembangan ilmu pengetahuan dan sebagai sumbangsih bagi masyarakat dalam membahas kegiatan vaksin *Covid-19*.

2. Manfaat Praktis

- a. Bagi Penulis

Dapat menerapkan algoritma *Naïve Bayes* dan seleksi fitur *Information Gain* untuk mendapatkan hasil akurasi optimal dalam klasifikasi opini dan menambah wawasan mengenai kegiatan vaksin *Covid-19*.

b. Bagi Masyarakat & Pemerintah

Manfaat dari hasil penelitian ini yaitu sebagai sarana motivasi masyarakat agar berpartisipasi dalam kegiatan vaksinasi *Covid-19* guna membantu pencegahan penularan *Covid-19*, yang mana hasil akhir analisis sentimen ini lebih dominan ke arah positif. Adapun jika hasil akhir dari analisis sentimen ini lebih dominan ke arah negatif maka akan membantu pemerintah penyelenggara kegiatan vaksin untuk dapat memikirkan serta menanggulangi dampak yang terjadi dikalangan masyarakat mengenai kegiatan vaksin *Covid-19*.

c. Bagi Universitas

Sebagai sumbangsih literatur kepustakaan tentang penelitian isu vaksin *Covid-19*.

1.6. Sistematika Penulisan

Sistematika dari penyusunan skripsi ini terdiri dari beberapa bagian yang paling utama untuk diketahui sebagai berikut:

BAB 1 PENDAHULUAN

Pada bab ini berisi tentang tahapan awal penelitian yang dimulai dari latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB 2 TINJAUAN PUSTAKA

Pada bab ini didalamnya akan dipaparkan mengenai penelitian-penelitian terdahulu dimana penelitian tersebut terhubung dengan permasalahan yang diteliti dan menjadi acuan konseptual.

BAB 3 METODOLOGI PENELITIAN

Pada bab ini akan dibahas mengenai teori dan konsep yang berhubungan dengan penelitian yang dilakukan untuk mendukung pemecahan masalah. Dan juga, bab ini memuat teori-teori dalam proses pengumpulan, pengolahan, dan penganalisisan data.

BAB 4 HASIL DAN PEMBAHASAN

Pada bab ini akan menjelaskan mengenai hasil pengumpulan, pengolahan dan analisis yang dilakukan.

BAB 5 KESIMPULAN DAN SARAN

Pada bab ini akan dipaparkan mengenai kesimpulan yang diperoleh dari hasil penelitian dan analisis data yang telah dilakukan juga saran-saran yang dapat diterapkan dari hasil pengolahan data agar dapat berguna.



DAFTAR PUSTAKA

- [1] W. Yulita, E. D. Nugroho, and M. H. Algifari, “Analisis Sentimen Terhadap Opini Masyarakat Tentang Vaksin Covid - 19 Menggunakan Algoritma Naïve Bayes Classifier,” vol. 2, no. 2, pp. 1–9, 2021.
- [2] “No Title.” <https://ourworldindata.org/explorers/coronavirus-data-explorer>.
- [3] F. Nurhuda, S. Widya Sihwi, and A. Doewes, “Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier,” *J. Teknol. Inf. ITSmart*, vol. 2, no. 2, p. 35, 2016, doi: 10.20961/its.v2i2.630.
- [4] S. Widya Sihwi, I. Prasetya Jati, and R. Anggrainingsih, “Twitter Sentiment Analysis of Movie Reviews Using Information Gain and Naïve Bayes Classifier,” *Proc. - 2018 Int. Semin. Appl. Technol. Inf. Commun. Creat. Technol. Hum. Life, iSemantic 2018*, pp. 190–195, 2018, doi: 10.1109/ISEMANTIC.2018.8549757.
- [5] E. D. Sri Mulyani, D. Rohpandi, and F. A. Rahman, “Analysis of Twitter Sentiment Using the Classification of Naive Bayes Method about Television in Indonesia,” *2019 1st Int. Conf. Cybern. Intell. Syst. ICORIS 2019*, vol. 1, no. August, pp. 89–93, 2019, doi: 10.1109/ICORIS.2019.8874896.
- [6] N. Ardhanie, R. Andreswari, and M. A. Hs, “Sentiment Analysis of ‘Indonesian No Dating Campaigns’ on Twitter Using Naïve Bayes Algorithm,” *Proc. - 2019 Int. Semin. Appl. Technol. Inf. Commun. Ind. 4.0 Retrospect. Prospect. Challenges, iSemantic 2019*, pp. 116–120, 2019, doi: 10.1109/ISEMANTIC.2019.8884331.
- [7] M. Arhami and M. Nasir, *Data Mining Argoritma dan Implementasi*. 2020.
- [8] S. Lestari and S. Saepudin, “Analisis Sentimen Vaksin Sinovac Pada Twitter Menggunakan Algoritma Naive Bayes,” *SISMATIK (Seminar Nas. Sist. Inf. dan Manaj. Inform.,* pp. 163–170, 2021.
- [9] Samsir, Ambiyar, U. Verawardina, F. Edi, and R. Watrianthos, “Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19

- Menggunakan Metode Naïve Bayes,” *J. Media Inform. Budidarma*, vol. 5, no. 1, p. 149, 2021, doi: 10.30865/mib.v5i1.2604.
- [10] S. yana nursyi’ah, “Analisis Sentimen Pembelajaran Daring Pada Masa Pandemi Covid-19 Di Twitter Menggunakan Algoritma Naive Bayes,” 2021.
- [11] S. S. Salim and J. Mayary, “Analisis Sentimen Pengguna Twitter Terhadap Dompot Elektronik Dengan Metode Lexicon Based Dan K – Nearest Neighbor,” *J. Ilm. Inform. Komput.*, vol. 25, no. 1, pp. 1–17, 2020, doi: 10.35760/ik.2020.v25i1.2411.
- [12] K. Sussolaikah, “Pemanfaatan Packages Pada R Programming Untuk Crawling Data Pada Social Media,” vol. 3, no. 3, pp. 203–206, 2021, doi: 10.47065/bits.v3i3.1035.
- [13] A. Turmudi and K. S. Yasah, “Analisa Sentimen Tweet Indonesia Menggunakan Fitur Ekstrasi Dan Teknik Cross Validation Terhadap Model Naive Bayes,” *SIGMA – J. Teknol. Pelita Bangsa*, vol. 10, pp. 189–194, 2020.
- [14] T. R. Suryanegara I.R.K, “Ekstrasi Informasi Dari Sampul Dokumen Skripsi,” 2019.
- [15] A. (UNIVERSITAS S. D. Febrianto, “Analisis Sentimen Terhadap Tokoh Publik Pada Twitter dengan Pendekatan Lexical Based,” 2020.
- [16] A. M. Hidayat, M. Syafrullah, and I. Pendahuluan, “Algoritma Naive Bayes Dalam Analisis Sentimen Untuk Klasifikasi Pada Layanan Internet PT . XYZ,” vol. 9, no. 2, pp. 91–95, 2017.
- [17] A. Syakuro, “Pada Media Sosial Menggunakan Metode Naïve Bayes Classifier (NBC) Dengan Seleksi Fitur Information Gain (IG) Halaman Judul Skripsi Oleh : Abdan Syakuro,” *Anal. sentimen Masy. terhadap e-commerce pada media Sos. menggunakan Metod. naive bayes Classif. dengan Sel. fitur Inf. gain*, pp. 1–89, 2017, [Online]. Available: <http://etheses.uin-malang.ac.id/11706/>.
- [18] A. B. P. Negara, H. Muhardi, and I. M. Putri, “Analisis Sentimen Maskapai

Penerbangan Menggunakan Metode Naive Bayes Dan Seleksi Fitur Information Gain Sentiment Analysis on Airlines Using Naïve Bayes Method and Feature Selection Information Gain,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 3, pp. 599–606, 2020, doi: 10.25126/jtiik.202071947.

- [19] I. A. Ridlo, *Panduan pembuatan flowchart*. 2017.
- [20] “Regular Expressions as used in R.” <https://stat.ethz.ch/R-manual/R-devel/library/base/html/regex.html>.

