

**ANALISIS SENTIMEN MASYARAKAT TERHADAP TINDAKAN  
VAKSINASI COVID-19 MENGGUNAKAN METODE NAÏVE BAYES**

Oleh  
**MELISNAWATI I. LAMARA**  
**T3118062**

**SKRIPSI**

Untuk memenuhi salah satu syarat ujian  
guna memperoleh gelar Sarjana



**PROGRAM SARJANA**  
**TEKNIK INFORMATIKA**  
**UNIVERSITAS ICHSAN GORONTALO**  
**GORONTALO**  
**2023**

**PERSETUJUAN SKRIPSI**

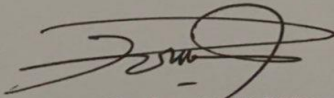
**ANALISIS SENTIMEN MASYARAKAT TERHADAP  
TINDAKAN VAKSINASI COVID-19 MENGGUNAKAN  
METODE NAÏVE BAYES**

Oleh  
**Melisnawati I. Lamara**  
T3118062

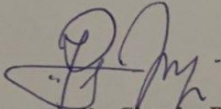
**SKRIPSI**

Untuk memenuhi salah satu syarat ujian  
guna memperoleh gelar Sarjana  
Program Studi Teknik Informatika,  
ini telah disetujui oleh Tim Pembimbing  
Gorontalo, Maret 2023

Pembimbing I

  
**Irvan Abraham Salih, M.Kom**  
NIDN : 0928028101

Pembimbing II

  
**Kartika Chandra Pelangi, M.Kom**  
NIDN : 0916038304

**PENGESAHAN SKRIPSI**

**ANALISIS SENTIMEN MASYARAKAT TERHADAP  
TINDAKAN VAKSINASI COVID-19 MENGGUNAKAN  
METODE NAÏVE BAYES**

Oleh  
Melisnawati I. Lamara  
T3118062

Diperiksa oleh Panitia Ujian Strata Satu (S1)  
Universitas Ichsan Gorontalo

1. Ketua Penguji  
Haditsah Annur, M.Kom
2. Anggota  
Sudirman Pana, M.Kom
3. Anggota  
Maryam Hasan, M.Kom
4. Anggota  
Irvan Abraham Salihi, M.Kom
- 5.. Anggota  
Kartika Chandra Pelangi, M.Kom

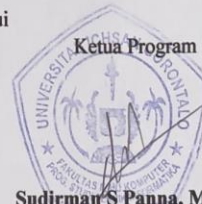
Mengetahui

Dekan Fakultas Ilmu Komputer



**Irvan Abraham Salihi, M.Kom**  
NIDN. 0928028101

Ketua Program Studi



**Sudirman S Panna, M.Kom**  
NIDN. 0924038205

## **ABSTRACT**

### **MELISNAWATI I. LAMARA. T3118062. ANALYSIS OF PUBLIC SENTIMENT TOWARDS COVID-19 VACCINATION MEASURES USING THE NAÏVE BAYES METHOD**

*Coronavirus is one of the most dangerous viruses due to its high and rapid spread rate throughout the world. World Health Organization (WHO) recorded many reports confirming coronavirus or Covid-19 infections from various countries in 2020. The spread of covid-19 is so fast and dangerous if no immediate action is taken, the way to prevent the spread of this virus is to develop a vaccine. Vaccines not only protect the vaccinated person but reduce the spread of disease in the public, especially in the Bongomeme subdistrict, Gorontalo District in Gorontalo Province. Covid-19 vaccination action can cause pro and con responses from the public. Therefore, a study is conducted which aims to determine the sentiment/response of the community to the vaccination action. This study is conducted with sentiment analysis using the Naïve Bayes Classifier method and weighting using the Term Frequency-Inverse Document Frequency (TF-IDF). The results of this study are the determination of the classification of public response questionnaire data so that the results are easy to read. The survey results from public responses to vaccination actions from 1,096 data contain 616 positive sentiments and 480 negative sentiments. The results obtained from the accuracy of the Naïve Bayes Classifier is 80%.*

*Keywords: Covid-19, sentiment analysis, Naïve Bayes Classifier, TF-IDF*

## ABSTRAK

### **MELISNAWATI I. LAMARA. T3118062. ANALISIS SENTIMEN MASYARAKAT TERHADAP TINDAKAN VAKSINASI COVID-19 MENGGUNAKAN METODE NAÏVE BAYES**

Coronavirus merupakan salah satu virus yang paling berbahaya, karena tingkat penyebarannya yang tinggi dan cepat menyebar keseluruh dunia. WHO mencatat banyaknya laporan yang mengonfirmasi infeksi virus corona atau Covid-19 dari berbagai negara pada tahun 2020. Penyebaran covid-19 yang begitu cepat dan berbahaya jika tidak segera diambil tindakan, cara untuk mencegah penyebaran virus ini adalah dengan mengembangkan vaksin. Vaksin tidak hanya melindungi orang yang divaksinasi tetapi dapat mengurangi penyebaran penyakit pada masyarakat umum, khususnya digorontalo kecamatan bongomeme kabupaten gorontalo bahwa tindakan vaksinasi Covid-19 dapat menimbulkan tanggapan pro dan kontra dari masyarakat oleh karena itu dilakukan penelitian yang bertujuan untuk mengetahui sentimen/tanggapan masyarakat terhadap tindakan vaksinasi. Penelitian ini dilakukan dengan analisis sentimen menggunakan metode Naïve Bayes Classifier dan pembobotan menggunakan Term Frequency-Invers Document Frequency (TF-IDF). Hasil penelitian ini untuk menentukan klasifikasi dari data kuisisioner tanggapan masyarakat sehingga hasilnya mudah dibaca. Hasil survei dari tanggapan masyarakat terhadap tindakan vaksinasi dari 1.096 data terdapat 616 sentimen positif dan 480 sentimen negatif. Hasil yang didapatkan dari akurasi Naïve Bayes Classifier memperoleh ketepatan 80%.

Kata kunci: Covid-19, analisis sentimen, Naïve Bayes Classifier, TF-IDF

## KATA PENGANTAR

Alhamdulillah, penulis dapat menyelesaikan skripsi ini dengan judul **“ANALISIS SENTIMEN MASYARAKAT TERHADAP TINDAKAN VAKSINASI COVID-19 MENGGUNAKAN METODE NAÏVE BAYES”**, untuk memenuhi salah satu syarat Ujian Akhir guna memperoleh gelar Sarjana Komputer pada Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.

Penulis menyadari sepenuhnya bahwa skripsi ini tidak mungkin terwujud tanpa bantuan dan dorongan dari berbagai pihak, baik bantuan moril maupun materil. Untuk itu, dengan segala keikhlasan dan kerendahan hati, penulis mengucapkan banyak terima kasih dan penghargaan yang setinggi-tingginya kepada:

1. Dr. Yuriko Abdusamat, M.si, selaku Ketua Yayasan Pengembangan Ilmu Pengetahuan dan Teknologi (YPIPT) Ichsan Gorontalo;
2. Dr. Abdul Gaffar La Tjokke, M.si, selaku Rektor Universitas Ichsan Gorontalo;
3. Irvan Abraham Salihi M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo sekaligus selaku Pembimbing I yang telah memberikan bimbingan dan arahan kepada penulis untuk menyelesaikan skripsi ini.
4. Sudirman Melangi, M.Kom, selaku Wakil Dekan I Bidang Akademik Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
5. Irma Surya Kumala Idris, M.Kom, selaku Wakil Dekan II Bidang Administrasi Umum dan Keuangan Fakultas Ilmu Komputer Universitas Ichsan
6. Sudirman S. Panna, M.Kom, selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
7. Kartika Chandra Pelangi, M.Kom, selaku Pembimbing II, yang telah memberikan bimbingan dan arahan kepada penulis untuk menyelesaikan skripsi ini;

8. Bapak dan Ibu Dosen Universitas Ichsan Gorontalo yang telah mendidik dan mengajarkan berbagai disiplin ilmu kepada penulis;
9. Teristimewa kepada keluarga saya, Ayah saya Ishak Lamara, Ibu saya Hasna Lasena, Kakak saya Desriwita I. Lamra, S.Ak, dan Okiandri I. Lamara, S.Gz terimakasih atas segala kasih sayang jerih payah, doa restu serta telah mendukung penulis mencapai cita-cita untuk menjadi seorang sarjana;
10. Rekan-rekan seperjuangan Fikom angkatan 2018
11. Kepada semua pihak yang telah membantu dalam penyelesaian skripsi ini tak sempat penulis sebutkan satu-persatu.
12. Terakhir, saya ingin mengucapkan terima kasih kepada diri saya sendiri karena sudah sabar, tetap bertahan, dan selalu sehat dalam melalui proses panjang perkuliahan Teknik Informatika hingga penyelesaian skripsi ini.

Semoga Allah SWT melimpahkan balasan atas jasa-jasa mereka kepada kami. Penulis menyadari sepenuhnya bahwa apa yang telah dicapai ini masih jauh dari kesempurnaan dan masih banyak terdapat kekurangan. Oleh karena itu, penulis sangat mengharapkan adanya kritik dan saran yang konstruktif. Akhirnya penulis berharap semoga hasil yang telah dicapai ini dapat bermanfaat bagi kita semua, Aamiin.

Gorontalo, Maret  
2023

Melisnawati I. Lamara

## DAFTAR ISI

<b>HALAMAN JUDUL .....</b>	<b>i</b>
<b>PERSETUJUAN SKRIPSI .....</b>	<b>ii</b>
<b>PENGESAHAN SKRIPSI.....</b>	<b>iii</b>
<b>PERNYATAAN SKRIPSI .....</b>	<b>iv</b>
<b>ABSTRACT .....</b>	<b>iv</b>
<b>ABSTRAK .....</b>	<b>v</b>
<b>KATA PENGANTAR.....</b>	<b>vi</b>
<b>DAFTAR ISI .....</b>	<b>viii</b>
<b>DAFTAR GAMBAR.....</b>	<b>xi</b>
<b>DAFTAR TABEL .....</b>	<b>xiii</b>
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1. 1. Latar Belakang.....	1
1. 2. Identifikasi Masalah .....	5
1. 3. Rumusan Masalah.....	5
1. 4. Tujuan Penelitian.....	5
1. 5. Manfaat Penelitian.....	5
<b>BAB II LANDASAN TEORI.....</b>	<b>7</b>
2. 1 Tinjauan Studi.....	7
2. 2 Tinjauan Pustaka.....	7
2.2.1 Covid-19.....	9
2.2.2 Vaksinasi Covid-19.....	10
2.2.3 Analisis Sentimen .....	10
2.2. <i>Text Pre- Preprocessing</i> .....	10



2.2.5 Parsing/Tokenizing .....	11
2.2.6 Filtering/Stopword.....	11
2.2.7 Stemming.....	11
2.2.8 Pembobotan TF-IDF .....	11
2.2.9 Proses Klasifikasin Naïve Bayes .....	12
2.2.10 Confusion Matrix.....	14
2.3 Kerangka Pikir.....	16
<b>BAB III METODE PENELITIAN.....</b>	<b>17</b>
3.1 Jenis, Metode, Subjek, Objek, Waktu dan Lokasi Penelitian.....	17
3.2 Pengumpulan Data.....	17
3.3 <i>Pre-Processing</i> .....	18
3.4 Pombobotan TF/IDF.....	18
3.5 Klasifikasi .....	19
3.7 Evaluasi .....	19
<b>BAB IV HASIL PENELITIAN .....</b>	<b>20</b>
4.1 Hasil Pengumpulan Data .....	20
4.1.1 Penerapan Metode .....	23
<b>BAB V HASIL DAN PEMBAHASAN.....</b>	<b>36</b>
5.1 Hasil Sistem.....	36
<b>BAB VI KESIMPULAN DAN SARAN .....</b>	<b>38</b>
6.1 Kesimpulan.....	38
6.2 Saran.....	38

## DAFTAR PUSTAKA

## LAMPIRAN

Hasil Turnitin

Listing Program

Dataset

Daftar Riwayat Hidup

Surat Penelitian

Surat Keterangan Penelitian

Surat Keterangan Bebas Plagiasi

Surat Keterangan Bebas Pustaka

## DAFTAR GAMBAR

Gambar 2.1 Kerangka Pikir.....	16
Gambar 3.1. Model yang diusulkan .....	18
gambar 5. 1 Confusion Matrix Without Normalization .....	35
Gambar 5. 2 confusion matrix normalized.....	36
Gambar 5.3 Hasil Akurasi.....	38

## DAFTAR TABEL

Tabel 1. 1: Jumlah Penduduk.....	1
Tabel 2. 1: Tinjauan Studi.....	3
Tabel 2.2 Confusion Matrix .....	14
Tabel 4. 1 Hasil Pengumpulan Data.....	21
Tabel 4. 2 Komentar Data Latih.....	24
Tabel 4. 3 <i>Case Folding</i> .....	24
Tabel 4. 4 <i>Tokenization</i> .....	25
Tabel 4. 5 <i>Filtering Stopword</i> .....	26
Tabel 4. 6 <i>Stemming</i> .....	26
Tabel 4. 7 Pembobotan TF.....	27
Tabel 4. 8 Hasil Pembobotan DF .....	27...
Tabel 4. 9 Hasil Pembobotan IDF.....	28
Tabel 4. 10 Pembobotan TF-IDF .....	29
Tabel 4. 11 Saran Data Uji.....	32
Tabel 4. 12 Data Uji Setelah <i>Preprocessing</i> .....	32
Tabel 4. 13 Tabel Data Latih dan Data Uji .....	32
Tabel 4. 14 Pembobotan Data Uji.....	33
Tabel 5.1 Hasil Confusion Matrix.....	36



# BAB I PENDAHULUAN

## 1.1 Latar Belakang

Corona Virus menjadi salah satu wabah dari November hingga Desember 2019 dikota Wuhan, China. virus ini merupakan salah satu virus yang paling berbahaya, karena tingkat penyebaran yang tinggi dan cepat menyebar ke seluruh dunia. Menurut catatn WHO, ada banyak laporan yang mengonfirmasi infeksi virus Corona atau Covid-19 dari berbagai negara ditahun 2020[1].

Mengingat penyebaran Covid-19 yang begitu cepat dan ada bahaya jika tidak segera diambil tindakan, cara yang paling mungkin untuk mencegah penyebaran virus ini adalah dengan mengembangkan vaksin. Vaksin tidak hanya melindungi orang yang divaksinasi, tetapi dengan mengurangi penyebaran penyakit, masyarakat umum juga [2].

Di Gorontalo khususnya Di Kecamatan Bongomeme Kabupaten Gorontalo bahwa tindakan vaksinasi COVID-19, dapat menimbulkan tanggapan pro dan kontra dari masyarakat. Meskipun tindakan vaksinasi tersebut diberikan secara gratis oleh pemerintah, agar tetapi sebagian besar masyarakat yang ada dikecamatan bongomeme tidak setuju untuk di vaksin.

**Tabel 1.1** Jumlah Penduduk Kecamatan Bongomeme

NO	DESA	JUMLAH PENDUDUK		
		L	P	L+P
1	DULAMYO	777	744	1.521
2	BONGOHULAWA	755	757	1.512
3	UPOMELA	1.178	1.144	2.322
4	TOHUPO	1.019	979	1.998
5	BATULORENG	435	554	989
6	MOLANIHU	635	440	875

NO	DESA	JUMLAH PENDUDUK		
		L	P	L+P
7	MOLOPOTADU	680	633	1313
8	MOLAS	856	797	1653
9	OTOPADE	617	592	1209
10	BATULAYAR	846	802	1648
11	HUNTULOHULAWA	509	559	1068
12	OWALANGA	466	413	879
13	LIYODU	348	344	692
14	LIYOTO	394	402	796
15	KAYUMERAH	546	489	1035
	JUMLAH	10.049	9.549	19.698

Berdasarkan hasil survei pemerintah kecamatan bongomeme yang memiliki target 15.000 ribu jiwa masyarakat, yang harus menerima tindakan vaksinasi covid-19. Akan tetapi ada beberapa masyarakat yang menolak untuk divaksin, Karena Masyarakat yang sudah melakukan vaksinasi dan yang belum divaksinasi memberikan tanggapan pro dan kontra di kehidupan sehari-hari.

Kebebasan masyarakat dalam mengutarakan pendapat menimbulkan berbagai opini baik positif maupun negatif. Agar vaksinasi dapat berjalan maksimal maupun berjalan lancar, pemerintah perlu mempertimbangkan beberapa masukan salah satunya dengan melihat tanggapan dari masyarakat terhadap pemberian vaksin COVID-19 tersebut. Perlu dilakukan pengolahan kata untuk mengevaluasi opini publik karena format data responnya tidak terstruktur, dengan tanda baca dan tidak ada bahasa default. Analisis sentimen digunakan untuk menyaring opini Publik dan mengklasifikasinya ke dalam kelas positif dan negatif. Sehingga hasil klasifikasi yang diperoleh dapat membantu pemerintah untuk mengetahui tanggapan dan kekhawatiran masyarakat terhadap vaksin Covid-19, memungkinkan pemerintah untuk mengevaluasi dan memutuskan strategi selanjutnya terkait dengan edukasi dan sosialisasi tentang vaksin Covid-19 kepada masyarakat.

Tabel 1.2 Sampel Data Kuisisioner Masyarakat

Nama	Alamat	Tanggapan
Isti Irawati Hanapi	Desa Liyodu	berbicara tentang covid,mnurut saya hal yg sekrng di perbincangkan oleh sejuta umat wkwkwkwk...bahkan di kenal sebagi virus yang sangat berbahaya.makanya di adakan vaksin nah afapun manfaat dari faksin, Tubuh seseorang yang telah disuntikkan vaksin, akan merangsang antibodi untuk belajar dan mengenali virus yang telah dilemahkan tersebut. Dengan demikian, tubuh akan mengenai virus dan mengurangi risiko terpapar.nah sya tidak setuju dengan adanya vaksin karena skrg saja yg sudah melakukan vaksin ke tahap 3 saja masi rentan terkena covid,yg biasa di kenal virus omicron.nah lantas apa gunanya di vaksin jika tubuh masi juga terpapar virus covid?
Aldealtri Slamet	Desa Liyodu	Menurut saya tindakan vaksinasi ini tidak efektif, saya sudah vaksin tapi tetap saja daya tahan tubuh masih saja rentan, kasus lain orang yang sudah vaksinasi tetap saja terkena covid, lantas untuk apa di lakukan vaksin, di vaksin maupun tidak di vaksin tetap



		saja akan terkena covid
Hatu Lasena	Desa Liyodu	saya tidak setuju di vaksin

*Sumber:* Google Form

tabel di atas menampilkan tanggapan masyarakat terhadap tindakan vaksinasi covid-19 saat ini. Berdasarkan tanggapan yang telah terkumpul maka akan diketahui tanggapan masyarakat terhadap tindakan vaksinasi covid-19. Salah satu untuk menentukan dan mengklasifikasi tanggapan *Positif* dan *Negatif* secara otomatis adalah dengan teknik analisis sentimen

Analisis sentimen, juga dikenal sebagai penggalian opini, adalah proses pemahaman, ekstrak dan proses data teks secara otomatis dapatkan informasi emosional yang terkandung dalam sebuah kalimat, apakah itu opini positif atau negatif. Analisis sentimen dapat menjadi solusi untuk masalah diatas. Dalam penelitian ini, Algoritma yang tepat untuk digunakan adalah Naïve Bayes Classifier.

Naïve Bayes Classifier digunakan untuk klasifikasi, artinya melihat pendapat dan kecenderungan seseorang tentang masalah dan objek, apakah kemungkinan akan diklasifikasikan sebagai negatif atau positif. Data yang telah mengalami proses pengolahan data teks melalui tahap klasifikasi menggunakan naïve bayes classifier untuk mengetahui apakah data tersebut dalam kategori positif atau negatif[3].

Berdasarkan penelitian sebelumnya yang telah dilakukan oleh “Dinda Ayu Muthia, 2017. Dengan judul penelitian “Analisis Sentimen Pada Review Restoran Dengan Teks Bahasa Indonesia Menggunakan Algoritma Naïve Bayes” Dari hasil data, genetic yang dilakukan algoritma terbukti meningkatkan akurasi pengklasifikasikan Naïve Bayes. Data peringkat restoran dapat diklasifikasikan dengan benar dalam forma Postif Dan Negatif. Akurasi Naïve Bayes mencapai 86,50% sebelum menggunakan pemilihan fitur, sisi lain setelah menggunakan kombinasi seleksi fitur, yaitu algoritma genetika, akurasinya ditingkatkan hingga 90,50%. Meningkatkan akurasi hingga 4%[4]. Maka dari pernyataan dan jurnal yang telah saya baca, sebagai peneliti akan menggunakan metode

naïve bayes karena metode tersebut sesuai dengan kebutuhan penelitian saya.

Berdasarkan uraian diatas, Maka penulis melakukan penelitian dengan judul **“Analisis Sentimen Masyarakat Terhadap Tindakan Vaksinasi Covid-19”**

## **1.2 Identifikasi Masalah**

Berdasarkan latar belakang masalah diatas, maka identifikasi masalah adalah kurangnya edukasi dan sosialisasi tentang vaksin Covid-19 kepada masyarakat. Sehingga kurangnya pemahaman masyarakat tentang vaksinasi covid-19

## **1.3 Rumusan Masalah**

Berdasarkan permasalahan yang telah di jelaskan di latar belakang, dapat diketahui rumusan masalah penelitian ini adalah **“Bagaimana Hasil Klasifikasi Data Teks Dalam Analisis Sentimen Masyarakat Terhadap Vaksin Covid-19 Di Kecamatan Bongomeme”**

## **1.4 Tujuan Penelitian**

Adapun Tujuan dari penelitian ini adalah Untuk mengetahui bagaiman hasil Analisis Sentimen Masyarakat Terhadap Vaksin Covid-19 Di Kecamatan Bongomeme”.

## **1.5 manfaat Penelitian**

Hasil penelitian ini diharapkan dapat memberikan manfaat antara lain sebagai berikut :

### **2.2.2 Manfaat Teoritis**

Berdasarkan hasil analisis yang dilakukan, hasil analisis sentimen tersebut diharapkan dapat digunakan oleh pemerintah sebagai tolak ukur dalam melakukan penilaian dan menyusun strategi edukasi dan sosialisasi Covid-19,

### **2.2.3 Manfaat Praktis**

Berdasarkan analisis sentimen yang dilakukan, diharapkan hasil analisis dapat menjadi sumber informasi serta pengetahuan yang menambah wawasan tentang pemahaman masyarakat terhadap vaksin Covid-19 sehingga pandemi Covid-19 dapat terkendalikan dan menurunkan angka kematian.

**BAB II**  
**LANDASAN TEORI**

**2.1 Tinjauan Studi**

**Tabel 2.1:** Tinjauan Studi

<b>NO</b>	<b>PENELITI</b>	<b>JUDUL</b>	<b>METODE</b>	<b>HASIL</b>
<b>1</b>	Dinda Ayu Muthia[4] (2017)	Analisis Sentimen Pada Review Restoran Dengan Teks Bahasa Indonesia Menggunakan Algoritma Naïve Bayes	Algoritma Naïve Bayes	Dari hasil data, genetic yang dilakukan algoritma terbukti meningkatkan akurasi pengklasifikasikan Naïve Bayes. Data peringkat restoran dapat diklasifikasikan dengan benar dalam forma Postif Dan Negatif. Akurasi Naïve Bayes mencapai 86,50% sebelum menggunakan pemilihan fitur, sisi lain setelah menggunakan kombinasi seleksi fitur, yaitu algoritma genetika, akurasinya ditingkatkan hingga 90,50%. Meningkatkan akurasi hingga 4%
<b>2</b>	-Adhi Viky Sudiantoro -Eri Zuliarso (2018)[3]	Analisis Sentimen Twitter Menggunakan Text Mining	Algoritma Naïve Bayes Classifier	berdasarkan hasil analisis dan pengujian yang telah dilakukan: 1. Algoritma klasifikasi naïve bayes sangat efektif digunakan sebagai

		Dengan Menggunakan Algoritman Naïve Bayes		<p>proses klasifikasi proses tweet yang dibutuhkan untuk analisis sentimen ini. Hasil yang dicapai dalam pengujian diperkirakan mencapai 84%</p> <p>2. Metode Naïve Bayes Classifier Sangat Baik Dalam Mengklasifikasikan Tweet Dalam Sistem Analisis Sentimen.</p> <p>3. 300 data yang dibagi menjadi 2, yaitu hingga 200 data pelatihan dan total 100 data pengujian. Hasil klasifikasi menunjukkan bahwa 100 data yang di uji termasuk dalam kategori emosi negatif.</p>
3	-Shima Fanissa -M. Ali Fauzi -Sigit Adinugroho (2018)[5]	Analisis Sentimen Parawisata Dikota Malang Menggunakan Metode Naïve Bayes Dan Seleksi Fitur Query Expansion Ranking	Algoritma Naïve Bayes	Metode <i>Multinomial Naïve Bayes</i> Dapat Diterapkan Pada Proses Analisis Sentimen Parawisata Malang Menggunakan Data Latih Dari Dari Ulasan-Ulasan Pada Web Tripadvisor. Data Latih Juga Harus diberi label dengan kelas suasana hati oleh ahli proses klasifikasi. Tahap dalam analisis sentimen adalah preprocessing(tokenisasi,

				<p>filtering, stemming), seleksi fitur Query Expansion Ranking, dan Naïve Bayes. Pemberian selekso fitur dalam proses analisis sentimen bertujuan untuk mereduksi dimensi fitur, selain itu metode peringkat yang diperluas query bekerja dengan baik dengan metode multinominal naïve bayes. Berdasarkan pengujian, algoritma peringkat ekstensi query mengembalikan akurasi tertinggi sebesar 86,6% dengan pemilihan fitur 75%.</p>
--	--	--	--	---

## 2.1 Tinjauan Pustaka

### 2.2.4 2.2.1 Covid-19

Coronavirus adalah sekelompok virus besar yang dapat menyebabkan penyakit dengan gejala ringan hingga berat. Penyakit Coronavirus 2019 (Covid-19) adalah jenis penyakit baru yang belum pernah ditemukan pada manusia sebelumnya[6].

Virus ini mulai masuk ke Indonesia sejak bulan maret 2020. Sejak virus corona ini ada di Indonesia telah mengalami banyak perubahan dalam kehidupan sehari-hari, dari kesehatan, budaya, sosial, transportasi, ekonomi, pendidikan dan lain-lain semuanya terdampak akibatvirus ini. Pemerintah telah mengendalikan penyebaran virus dengan berbagai cara, termasuk Pemberlakuan Pembatasan Sosial Berskal Besar (PSBB). Pemerintah juga

mengadakan sosialisasi seperti memakai masker, mencuci tangan, dan menjaga jarak sosial (3 M). [7]

### **2.2.5 Vaksinasi Covid-19**

Vaksin adalah produk biologis yang mengandung antigen yang aktif ketika diberikan kepada manusia, kembangkan kekebalan khusus penyakit tertentu. Berbagai negara termasuk Indonesia, pengembangan vaksin yang sangat efektif untuk mencegah infeksi SARS-CoV-2 di berbagai platform, yaitu vaksin virus yang dilemahkan, vaksin hidup di lemahkan, vaksin vektor virus, vaksin asam nukleat, seperti virus.(vaksin mirip virus) dan vaksin subunit protein. Tujuan pembuatan vaksin adalah untuk mengendalikan penyebaran Covid-19, mengurangi angka kesakitan dan kematian akibat Covid-19, mencapai imunitas kelompok dan melindungi masyarakat dari Covid-19, sehingga menjaga produktivitas sosial dan ekonomi.[6]

### **2.2.6 Analisis Sentimen**

Analisis sentimen digunakan untuk memahami dan mengolah data yang tekstual dan memperoleh informasi dalam pernyataan opini. Tujuan dari analisis sentimen adalah untuk mengidentifikasi pendapat (positif, negatif, atau netral) atau kecenderungan pendapat tentang masalah atau tujuan seseorang.[8]

Mekanisme analisis sentimen adalah dengan menggunakan pembelajaran mesin *machine learning* sebagai alat untuk membuat “robot” yang dapat digunakan untuk mengklasifikasikan untuk tipe emosi.[9]

Langkah pertama membutuhkan data teks (kalimat opini) yang sudah di beri label oleh annotator yaitu orang bertanggung jawab dalam pemberian label positif, negatif dan netral pada masing-masing data [10]

### **2.2.7 Text Pre-Processing**

Fungsi utama dari *Pre-processing teks* adalah untuk membersihkan data, hilangkan atau atasi data yang bising dan temukan informasi yang

tidak lengkap. Tujuan dari tahap ini adalah untuk mengoptimalkan hasil perhitungan.[11]

### 2.2.8 Parsing/Tokenizing

Tahap ini membantu setiap kalimat menjadi kata/token dalam dokumen. Secara umum pemisahan kata dengan spasi, Namun tujuannya tetap sama, sehingga dapat menggunakan fungsi pemisah apapun. Artinya memotong kalimat berdasarkan kata-kata yang membentuk kalimat.[11]

### 2.2.9 Filtering/Stopword Removal

Pada tahap ini , tujuannya adalah untuk menghilangkan kata-kata yang tidak penting, seperti kata penghubung dan kata ganti. Stoplist/Stopwords adalah kata-kata tidak berarti yang dapat dibuang dengan pendekatan bag-of-words[12].

### 2.2.10 Stemming

Stemming digunakan untuk mengonversi kata dalam dokumen ke format kata dasar. Dalam tahap stemming pada bahasa indonesia, perlu menghapus awalan kata dan tambahan akhir dalam dokumen [11].

### 2.2.11 Pembobotan TF-IDF

*Term Frequency-Inverse Document Frequency*(TF-IDF) adalah proses untuk ekstraksi fitur dengan cara memberikan bobot nilai pada masing-masing kata yang terdapat pada dokumen, atau bisa dikatakan mengubah kata menjadi bentuk numerik. Fungsi dari tahap ini adalah untuk menghitung seberapa penting sebuah kata dalam suatu dokumen cara menentukan nilai TF-IDF adalah seperti dibawah ini:

Langkah untuk melakukan pembobotan TF-IDF adalah:[11]

1. Menentukan kata kunci
  - a. Menentukan nilai TF bisa menggunakan biner, yaitu jika suatu kata terdapat dalam sebuah dokumen akan diberi nilai 1(satu), jika tidak akan diberi nilai 0(nol),
  - b. Menentukan nilai  $df$  dengan cara menghitung jumlah dokumen yang mendukung kata tertentu.



- c. Hitung nilai frekuensi dokumen terbalik (IDF) yang berguna untuk menentukan pentingnya kata-kata dalam dokumen. IDF dirumuskan sebagai berikut.

$$\text{IDF} = \log \frac{D}{df} \quad (1)$$

Keterangan:

$df$  = jumlah dokumen terhitung yang mengandung istilah

$D$  = jumlah total dokumen

- d. Setelah menghitung TF-IDF digabungkan perhitungan TF dengan IDF sebagai berikut.

$$W_{ij} = tf \times \text{IDF} \quad (2)$$

$$W_{ij} = tf \times \log \frac{D}{df} \quad (3)$$

Keterangan:

$W_{ij}$  = bobot kata dalam setiap dokumen

$tf$  = jumlah kemunculan kata dalam dokumen

$D$  = jumlah seluruh dokumen

### 2.2.12 Proses Klasifikasi Naïve Naves

Tujuan dari proses klasifikasi adalah untuk menentukan apakah kalimat termasuk kedalam kelas opini positif dan negatif atau sebagai anggota kelas opini netral, yang ditentukan berdasarkan perhitungan probabilitas Bayesian yang lebih besar, yaitu. Jika probabilitas bayesian suatu kalimat pada kelas opinini positif memiliki hasil probablitas bayesian yang besar, maka kalimat tersebut ditempatkan pada kategori opini positif dan sebaliknya[12].

Algoritma Naïve Bayes merupakan salah satu algoritma yang termasuk dalam klasifikasi. Naïve bayes merupakan fungsi statistik sederhana berdasarkan teorema bayes dan mengasumsikan bahwa properti tertentu ada dikelas yang tidak bergantung pada properti lainnya.[14]

Persamaan dari kemungkinan:

$$P(H) = \frac{N_j}{N}$$

Keterangan :

$N_j$  : Nomor Data Kelas

$N$  : Total Data

Dimana

Persamaan Teorema Bayes:

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

Keterangan :

$X$  : Belum Diketahui Data Kelas

$H$  : Hipotesis data kelas spesifik

$P(H|X)$  : Kondisi  $X$  probabilitas hipotesis  $H$

$P(H)$  : Probabilitas  $X$  dalam keadaan Hipotesis  $H$

$P(X)$  : Probabilitas  $H$

Persamaan kedua Metode *Naïve Bayes*:

$$P(H|F_1, \dots, F_n) = \frac{P(C)P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)}$$

Pada persamaan (2) variabel  $C$  adalah sebuah kelas, sedangkan  $F_1 \dots F_n$  adalah karakteristik dari petunjuk yang di perlukan dalam proses klasifikasi.

Persamaan ketiga Metode *Naïve Bayes*:

$$\text{Posterior} = \frac{\text{prior} \times \text{likelihood}}{\text{Evidence}}$$

Dimana nilai evidence selalu sama untuk setiap kelas dalam sampel. Nilai posterior kemudian dibandingkan dengan nilai kelas posterior dalam kategori sampel yang akan diklasifikasikan [].

Persamaan keempat Metode *Naïve Bayes*:

$$\begin{aligned} P(C|X_1, \dots, X_n) &= P(H) P(X_1, \dots, X_n|C) \\ &= P(H)P(X_1|H) P(X_2, \dots, X_n|H, X_1) \\ &= P(H)P(X_1|H) P(X_2|H, X_1)P(X_3, \dots, X_n|H, X_1, X_2) \\ &= P(H)P(X_1|H)P(X_2|H, X_1)P(X_3|H, X_2, X_2) \\ &\dots P(X_n|H, X_1, X_2, \dots, X_{n-1}) \end{aligned}$$

Persamaan kelima Metode *Naïve Bayes*:

$$P(w|\text{positif/negatif}) = \frac{\text{count}(\text{training}_{\text{set}}, \text{sentimen} + 1)}{\text{count}(\text{sentimen}) + |v|}$$

### 2.2.10 Confusion Matrix

*Confusionmatrix* digunakan sebagai metode untuk mengevaluasi keefektifan atau kinerja model klasifikasi yang dibuat melalui pengukuran dan pengujian. Variabel uji yang di gunakan dalam proses evaluasi adalah akurasi, tabel yang di tampilkan dalam confusion matrix terdiri dari kelas predikat dan kelas sebenarnya. Selama proses pengujian, model menerima hasil nilai matriks akurasi dan konfusi[13].

**Tabel 2.2** *Confusion Matrix*

confusion matrix		Kelas Aktual	
		Positif	Negatif
kelas prediksi	Positif	TP	FP
	Negatif	FN	TN

Keterangan dari nilai matriks adalah:

- True Positive* (TP) adalah sejumlah data yang diprediksi positif dan kenyataan positif.
- True Negative* (TN), adalah sejumlah data yang diprediksi negatif dan kenyataannya negatif.
- False Positive* (FP), adalah sejumlah data yang diprediksi positif dan kenyataan negatif.
- False Negative* (FN), adalah sejumlah data yang diprediksi negatif dan kenyataan positif

Pada *confusion matrix* ada beberapa parameter yang digunakan untuk mengukur performa suatu metode yaitu[14]:

- Accuracy merupakan perbandingan kasus yang diidentifikasi dengan benar dengan jumlah total kasus

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN}$$

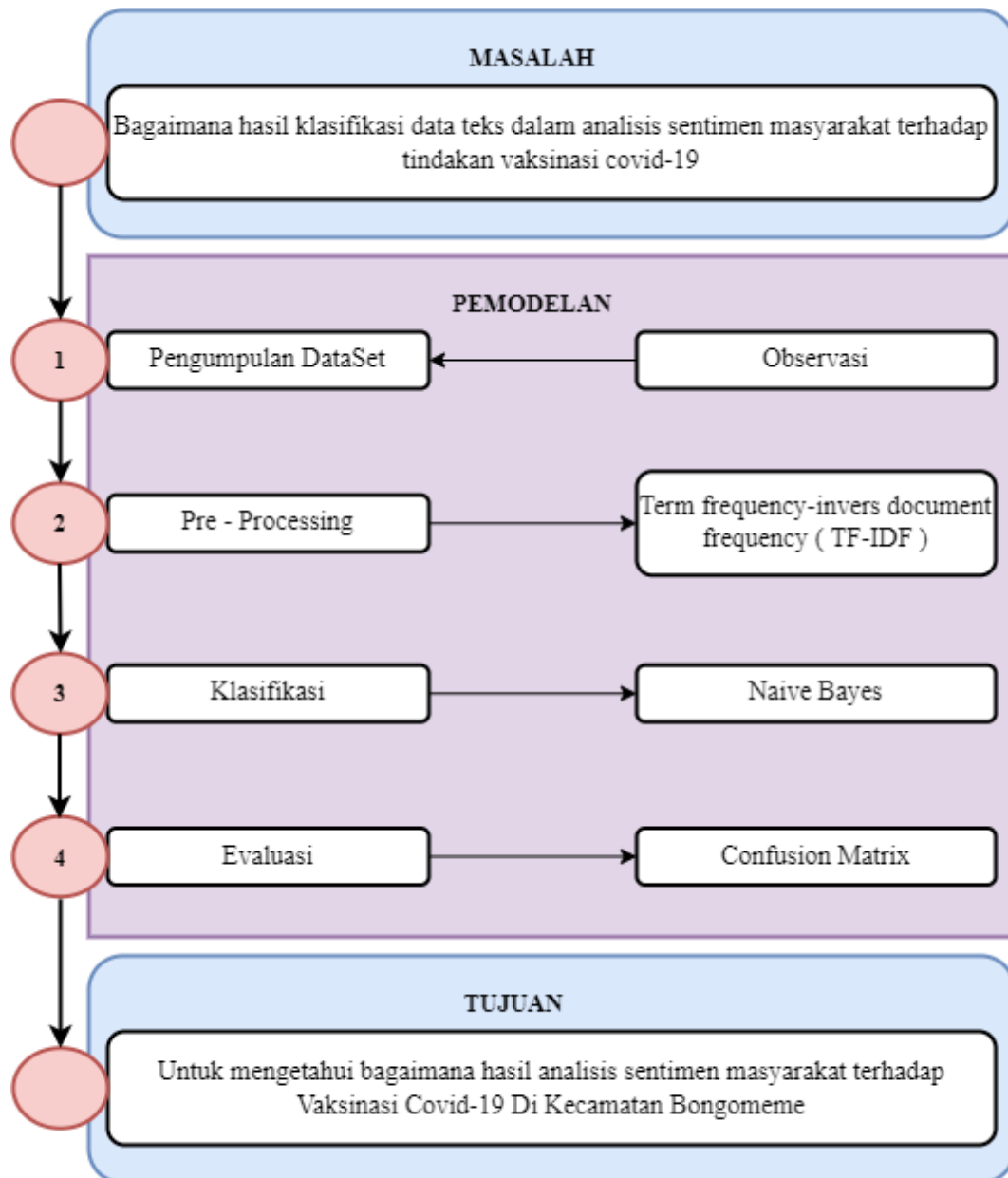
- Precesion* adalah sebagian kecil dari kasus positif

$$\frac{TP}{TP + FP}$$

c. *Recal* adalah proporsi kasus positif yang teridentifikasi dengan benar.

$$\frac{TP}{FN + TP}$$

## 2.1 Kerangka Pikir



**gambar 2.1:** Kerangka Pikir

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Jenis, Metode, Subjek, Objek, Waktu, dan Lokasi Penelitian**

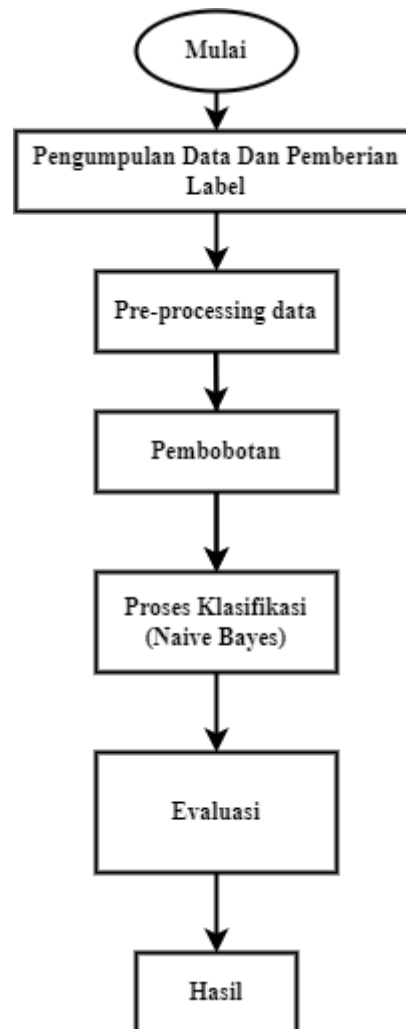
Dilihat dari tingkatannya, penelitian ini termasuk penelitian terapan. Dilihat dari jenis data yang diolah, penelitian ini adalah penelitian kuantitatif. Dilihat dari jenis data, maka penelitian ini merupakan penelitian konfirmatori.

Penelitian ini menggunakan metode eksperimen, dengan demikian jenis penelitian ini adalah menggunakan metode eksperimen, subjek dari penelitian ini adalah analisis sentimen masyarakat terhadap tindakan vaksinasi covid-19. Penelitian ini dimulai dari Januari 2022 s/d Februari 2022, yang berlokasi di Kecamatan Bongomeme, Kabupaten Gorontalo.

#### **3.2 Pengumpulan Data**

Data primer ini merupakan data hasil kuesioner masyarakat yang ada di Kecamatan Bongomeme yang dikumpulkan menggunakan teknik Kuesioner. Terdapat jumlah total data “1096” tanggapan masyarakat. Data yang akan diolah terdapat pada kolom tanggapan.

Dan untuk data sekunder merupakan data yang sudah ada diberbagai sumber referensi seperti buku, jurnal dan hasil penelitian terdahulu yang berkaitan dengan penelitian terkait yaitu analisis sentimen masyarakat terhadap tindakan vaksinasi Covid-19.



Gambar 3. 1 Model yang diusulkan

### 3.2 Pre-Processing

Data yang dikumpulkan dan dibagi menjadi data latih dan data uji melalui tahap pre-processing menggunakan Python. Proses ini membantu untuk mengklasifikasikan data dan mempermudah proses analisis.[1]

### 3.3 Pembobotan TF/IDF

Langkah Selanjutnya adalah memberi bobot pada setiap kata berdasarkan frekuensi dari term atau istilah yang muncul pada dokumen. Pembobotan ini dilakukan dengan metode TF/IDF [1].

### **3.4 Klasifikasi**

Proses klasifikasi dilakukan menggunakan algoritma naïve bayes. Proses ini dilakukan setelah data yang digunakan melalui proses pembobotan [1].

### **3.5 Evaluasi**

Evaluasi dilakukan untuk memverifikasi hasil klasifikasi dengan mengukur keakrutan dari suatu metode. Tahap evaluasi dilakukan dengan menggunakan metode *confussion matrix* untuk membandingkan hasil klasifikasi yang dilakukan oleh sistem sesuai dengan hasil yang sebenarnya.



## **BAB IV**

### **HASIL PENELITIAN**

#### **4.1 Hasil Pengumpulan Data**

Data yang digunakan dalam penelitian ini merupakan data tanggapan masyarakat terhadap tindakan vaksinasi sebanyak 1096 data. Data ini dibagi menjadi 2 (dua) data yaitu data latih dan data uji, dimana data latih sebanyak 767 data dan data uji sebanyak 329 data.

Data tanggapan masyarakat terhadap tindakan vaksinasi ini diambil secara online lewat *google form* kemudian disalin ke *file excel* untuk dimasukkan kedalam database. Data file excel ini diseleksi dan dilabeli secara manual terlebih dahulu untuk menghapus data-data yang kosong dan tidak sesuai diproses. Selanjutnya data tanggapan yang ada pada database dilakukan preprocessing dan disimpan dalam database data bersih.

Berdasarkan hasil pengumpulan data, diperoleh data primer sebagai berikut :

**Tabel 4.1 Hasil Pengumpulan Data**

No	Nama	Alamat	Komentar	Kelas
1	isti irmawati hanapi	desa liodu	berbicara tentang covid,mnurut saya hal yg sekrng di perbincangkan oleh sejuta umat wkwkwkwk...bahkan di kenal sebagi virus yang sangat berbahaya.makanya di adakan vaksin nah afapun manfaat dari faksin, Tubuh seseorang yang telah disuntikkan vaksin, akan merangsang antibodi untuk belajar dan mengenali virus yang telah dilemahkan tersebut. Dengan demikian, tubuh akan mengenai virus dan mengurangi risiko terpapar.nah sya tidak setuju dengan adanya vaksin karena skrg saja yg sudah melakukan vaksin ke tahap 3 saja masi rentan terkenah covid,yg biasa di kenal virus omicron.nah lantas apa gunanya di vaksin jika tubuh masi juga terpapar virus covid?	1
2	Wilan	Desa liodu	Sangat meresahkan masyarakat karena dalam mengurus data-data penting harus sudah	0

No	Nama	Alamat	Komentar	Kelas
			melakukan vaksinasi	
3	Yeyen Hemeto	Desa Liyodu	Menurut saya tindakan vaksinasi ini tidak efektif, saya sudah vaksin tapi tetap saja daya tahan tubuh masih saja rentan, kasus lain orang yang sudah vaksinasi tetap saja terkena covid, lantas untuk apa di lakukan vaksin, di vaksin maupun tidak di vaksin tetap saja akan terkena covid	0
4	Rolan ma'ruf	Desa liyodu kec. Bongomeme	vaksin tidak halal	0
5	Maryam Lasena	Desa liyodu	Sebagian masyarakat berpendapat terjadi pemaksaan dalam vaksinasi saat ini	0
6	Saida K. Dukalang	Desa Liyodu, Kec.Bongomeme	saya tidak setuju di vaksin	0
7	Sutika	Desa Liyodu	Pendapat saya tentang vaksinasi tindakan ini merupakan percobaan kepada tubuh manusia terhadap bakteri yg di diberikan kedalam tubuh, hal ini menurut saya tidaklah berpengaruh terhadap kebugaran tubuh karena hal ini membuat tubuh manusia menjadi rentan terhadap bakteri	1

No	Nama	Alamat	Komentar	Kelas
			yang di berikan kedalam tubuh manusia, dengan hal ini saya kurany setuju dengan adanya vaksinasi ini terlebih vaksinasi digunakan untuk administrasi yang tidak jelas.	
8	Hadijah lasena	Desa liyodu	Bisa melindungi saya dari paparan covid -19	1
9	Hatu k.lasena	Desa liyodu	Saya masih ragu untuk divaksin karna banyak yang mengatakan ada efek samping	0
10	Yunus Lasena	Desa liyodu	Saya bersyukur dengan adanya vaksin ini karna dengan vaksin ini bisa menjaga tubuh saya dari virus	1
...	...	....	.....	.....
1097	sri fani G lamara	desa molanihu	Vaksin Covid 19 Itu Wajib Karna Menjegah Penyakit Yang Akan Datang	1

Keterangan :

Tanggapan Positif : 1

Tanggapan Negatif : 0

#### 4.1.1 Penerapan Metode pada Data latih

Berikut ini merupakan perhitungan manual penyelesaian menggunakan metode *Naïve Bayes Classifier* dalam klasifikasi komentar.

Tabel 4.2 Komentar Data Latih

No	Nama	Alamat	Komentar	Kelas
1	Narti Musa	Desa Bongohulawa	Vaksinasi covid-19 saat ini membuat masyarakat risih karena sebagian masyarakat selesai melakukan vaksinasi mengalami sakit ( demam flu dan batuk )	0
2	Rio Ismail	Desa tohupo	Vaksin aman	1

1. Tahap *Preprocessing* :

- ✚ *Case Folding* : merubah huruf besar menjadi huruf kecil dan menghilangkan seluruh tanda baca.

Tabel 4.3 *Case Folding*

No	Komentar	Hasil <i>Case Folding</i>
1	Vaksinasi covid-19 saat ini membuat masyarakat risih karena sebagian masyarakat selesai melakukan vaksinasi mengalami sakit ( demam flu dan batuk )	vaksinasi covid-19 saat ini membuat masyarakat risih karena sebagian masyarakat selesai melakukan vaksinasi mengalami sakit ( demam flu dan batuk )
2	Vaksin aman	vaksin aman

- ✚ *Tokenization* : Pemecahan Teks/Kalimat menjadi kumpulan kata agar mudah dalam melakukan pembobotan tiap kata.

**Tabel 4.4 Tokenization**

Data 1	Data 2
vaksin	vaksinasi
aman	Covid
	masyarakat
	Risih
	selesai
	alami
	Sakit
	demam
	Flu
	Batu

+ *Filtering/stopword* : tahap untuk menghilangkan data dari karakter-karakter atau kata yang tidak berguna

**Tabel 4.5 Filtering/Stopword**

Sebelum Filtering/Stopword	Sesudah Filtering/Stopword
vaksin	vaksin
aman	aman
vaksinasi	vaksinasi
covid-19	covid
saat	
ini	
membuat	
masyarakat	masyarakat
risih	risih
karena	
sebagian	
masyarakat	

Sebelum Filtering/Stopword	Sesudah Filtering/Stopword
selesai	selesai
melakukan	
vaksinasi	
mengalami	alami
sakit	sakit
demam	demam
flu	flu
dan	
batuk	batuk

✚ *Stemming* : merubah kata berimbuhan menjadi kata dasar.

**Tabel 4.6 *Stemming***

Sebelum <i>Stemming</i>	Hasil <i>Stemming</i>
vaksinasi	vaksinasi
covid	covid
masyarakat	masyarakat
risih	risih
selesai	selesai
mengalami	alami
sakit	sakit
demam	demam
flu	flu
batuk	batuk
vaksin	vaksin
aman	aman

2. Pembobotan : menggunakan *tf(term frequency)* untuk mengetahui beberapa sering suatu kata muncul didalam dokumen.

- ❖ Langkah pertama yang dilakukan menghitung TF dengan cara menghitung tiap bobot term dalam teks. Dengan memberikan bobot nilai 1 dan 0.

**Tabel 4.7 Pembobotan TF**

No	Kosa Kata	Tf Positif	Tf Negatif
1	vaksin	1	0
2	aman	1	0
3	vaksinasi	0	1
4	covid	0	1
5	masyarakat	0	1
6	risih	0	1
7	selesai	0	1
8	alami	0	1
9	sakit	0	1
10	demam	0	1
11	flu	0	1
12	batuk	0	1
Jumlah Term		2	10

Diperoleh :

Count positif = 2, count negative = 10

Dengan Total jumlah Kata = 12 kata

- ❖ Langkah kedua menghitung DF dengan menjumlah nilai bobot nilai term dalam Dokumen.

**Tabel 4.8 Tabel Hasil Pembobotan DF**

No	Term	TF		DF
		K1	K2	
1	vaksin	1	0	1
2	aman	1	0	1
3	vaksinasi	0	1	1
4	covid	0	1	1



No	Term	TF		DF
		K1	K2	
5	masyarakat	0	1	1
6	risih	0	1	1
7	selesai	0	1	1
8	alami	0	1	1
9	sakit	0	1	1
10	demam	0	1	1
11	flu	0	1	1
12	batuk	0	1	1

❖ Langkah ketiga Menghitung IDF dengan menggunakan rumus :

$$\text{IDF} : \log(D/F)$$

Dimana :

D : Total dokumen

Df : banyak dokumen yang mengandung kata yang dicari

**Tabel 4.9 Hasil Pembobotan IDF**

No	Term	TF		DF	IDF
		D1	D2		
1	vaksin	1	0	1	0,301029996
2	aman	1	0	1	0,301029996
3	vaksinasi	0	1	1	0,301029996
4	covid	0	1	1	0,301029996
5	masyarakat	0	1	1	0,301029996
6	risih	0	1	1	0,301029996
7	selesai	0	1	1	0,301029996
8	alami	0	1	1	0,301029996
9	sakit	0	1	1	0,301029996
10	demam	0	1	1	0,301029996

11	flu	0	1	1	0,301029996
12	batuk	0	1	1	0,301029996

❖ Langkah Keempat menghitung TF-IDF

Pada tabel 4.15 adalah hasil pembobotan berdasarkan data di tabel sebelumnya. Bobot akhir suatu term mengalikan keduanya yaitu **TF x IDF**.

**Tabel 4.10 Pembobotan TF-IDF**

No	Term	TF		DF	IDF	TF x IDF	
		D1	D2			D1	D2
1	vaksin	1	0	1	0,301029996	0,301029996	0
2	aman	1	0	1	0,301029996	0,301029996	0
3	vaksinasi	0	1	1	0,301029996	0	0,301029996
4	covid	0	1	1	0,301029996	0	0,301029996
5	masyarakat	0	1	1	0,301029996	0	0,301029996
6	risih	0	1	1	0,301029996	0	0,301029996
7	selesai	0	1	1	0,301029996	0	0,301029996
8	alami	0	1	1	0,301029996	0	0,301029996
9	sakit	0	1	1	0,301029996	0	0,301029996
10	demam	0	1	1	0,301029996	0	0,301029996
11	flu	0	1	1	0,301029996	0	0,301029996
12	batuk	0	1	1	0,301029996	0	0,301029996

#### 4.1.2 Proses Klasifikasi Metode Naïve Bayes Pada Data Latih

1. Menghitung probabilitas *prior* dari setiap kategori, yaitu menjadi 2 kategori yakni kelas positif dan kelas negatif

$$P(\text{positif}) = \frac{n(\text{positif})}{n} = \frac{1}{2} = 0,5$$

$$P(\text{negatif}) = \frac{n(\text{negatif})}{n} = \frac{1}{2} = 0,5$$

2. Menghitung probabilitas *likelihood* atau mencari nilai peluang setiap term dari dokumen. Jumlah seluruh kata 12, 2 term dari kelas positif dan 10 term dari kelas negatif. Banyaknya term tergantung pada hasil *prerprocessing* data.

$$P(w|\text{positif/negatif}) = \frac{\text{count}(\text{training}_{\text{set}}, \text{sentimen} + 1)}{\text{count}(\text{sentimen}) + |v|}$$

1. Probabilitas kata “ vaksin”

$$P(\text{vaksin}|\text{positif}) = \frac{1 + 1}{2 + 12} = 0,1428571429$$

$$P(\text{vaksin}|\text{negatif}) = \frac{0 + 1}{10 + 12} = 0,0454545455$$

2. Probabilitas kata “ aman “

$$P(\text{aman}|\text{positif}) = \frac{1 + 1}{2 + 12} = 0,1428571429$$

$$P(\text{aman}|\text{negatif}) = \frac{0 + 1}{10 + 12} = 0,0454545455$$

3. Probabilitas kata “vaksinasi”

$$P(\text{vaksinasi}|\text{positif}) = \frac{0 + 1}{2 + 12} = 0,0714285714$$

$$P(\text{vaksinasi}|\text{Negatif}) = \frac{1 + 1}{10 + 12} = 0,0909090909$$

4. Probabilitas kata “ covid”

$$P(\text{covid}|\text{positif}) = \frac{0 + 1}{2 + 12} = 0,0714285714$$

$$P(\text{covid}|\text{negatif}) = \frac{1 + 1}{10 + 12} = 0,0909090909$$

5. Probabilitas kata “ Masyarakat”

$$P(\text{Masyarakat}|\text{positif}) = \frac{0 + 1}{2 + 12} = 0,0714285714$$

$$P(\text{Masyarakat}|\text{negatif}) = \frac{1 + 1}{10 + 12} = 0,0909090909$$

6. Probabilitas kata “ risih”

$$P(\text{risih}|\text{positif}) = \frac{0 + 1}{2 + 12} = 0,0714285714$$

$$P(\text{risih}|\text{negatif}) = \frac{1 + 1}{10 + 12} = 0,0909090909$$

7. Probabilitas kata “ selesai “

$$P(\text{Selesai}|\text{positif}) = \frac{0 + 1}{2 + 12} = 0,0714285714$$

$$P(\text{selesai}|\text{negatif}) = \frac{1 + 1}{10 + 12} = 0,0909090909$$

8. Probabilitas kata “ alami”

$$P(\text{Selesai}|\text{positif}) = \frac{0 + 1}{2 + 12} = 0,0714285714$$

$$P(\text{selesai}|\text{negatif}) = \frac{1 + 1}{10 + 12} = 0,0909090909$$

9. Probabilitas kata “ sakit”

$$P(\text{Sakit}|\text{positif}) = \frac{0 + 1}{2 + 12} = 0,0714285714$$

$$P(\text{Sakit}|\text{negatif}) = \frac{0 + 1}{10 + 12} = 0,0909090909$$

10. Probabilitas kata “ demam”

$$P(\text{demmt}|\text{positif}) = \frac{0 + 1}{2 + 12} = 0,0714285714$$

$$P(\text{demam}|\text{negatif}) = \frac{1 + 1}{10 + 12} = 0,0909090909$$

11. Probabilitas kata “ flu”

$$P(\text{flu}|\text{positif}) = \frac{0 + 1}{2 + 12} = 0,0714285714$$

$$P(\text{flu}|\text{negatif}) = \frac{1 + 1}{10 + 12} = 0,0909090909$$

12. Probabilitas kata “ batuk”

$$P(\text{Batuk}|\text{positif}) = \frac{0 + 1}{2 + 12} = 0,0714285714$$

$$P(\text{Batuk}|\text{negatif}) = \frac{1 + 1}{10 + 12} = 0,0909090909$$

**Tabel 4.11 Saran Data Uji**

No	Nama	Alamat	Tanggapan	value
1	Elen hubulo	Desa liyodu	Vaksin aman	1
2	Rastra pratiwy tolinggi	Desa liyodu	Sangat bagus untuk masyarakat menghindari covid yg tersebar saat ini	1

**Tabel 4.12 Data Uji Setelah Preprocessing**

No	Nama	Hasil Preprocessing	Value
1	Elen hubulo	Vaksin aman	1
2	Rastra pratiwy tolinggi	bagus masyarakat hindar covid sebar	1

**Tabel 4.13 Tabel Data Latih Dan Data Uji**

No	Nama	Hasil Preprocessing	Value
1	Narti Musa	vaksinasi covid masyarakat risih selesai alami sakit demam flu batuk	0
2	Rio Ismail	Vaksin aman	1
Data Uji			
1	Elen hubulo	Vaksin aman	?
2	Rastra pratiwy tolinggi	bagus masyarakat hindar covid sebar	?

#### 4.1.3 Proses Klasifikasi Naïve Bayes pada Data Uji

Proses klasifikasi data uji yaitu dengan mengalikan semua nilai peluang. Nilai yang lebih tinggi merupakan kelas dari data tersebut.

1. Pada data uji vaksin aman dan halal yang termasuk dalam dalam data latih “vaksin” dan “aman”

$$P(\text{uji}|\text{positif})$$

$$= P(\text{positif}) \times P(\text{vaksin}|\text{positif}) \times P(\text{aman}|\text{positif})$$

$$= 0,5 \times 0,1428571429 \times 0,1428571429$$

$$= 0,0102040817$$

$$P(\text{uji}|\text{negatif})$$

$$= P(\text{negatif}) \times P(\text{vaksin}|\text{negatif}) \times P(\text{aman}|\text{negatif})$$

$$= 0,5 \times 0,0454545455 \times 0,0454545455$$

$$= 0,0010330579$$

Kesimpulan : Nilai probabilitas tertinggi yaitu sebesar 0,0010330579 pada  $P(\text{uji}|\text{positif})$  sehingga komentar tersebut diklasifikasikan ke dalam kelas “**Positif**”

2. Pada data uji “bagus masyarakat hindar covid sebar” yang termasuk dalam data latih adalahb ” masyarakat”

$$P(\text{uji}|\text{positif})$$

$$= P(\text{positif}) \times P(\text{masyarakat}|\text{positif})$$

$$= 0,5 \times 0,0714285714$$

$$= 0,0357142857$$

$$P(\text{uji}|\text{negatif})$$

$$= P(\text{uji}|\text{negatif}) \times P(\text{masyarakat}|\text{negatif})$$

$$= 0,5 \times 0,0909090909$$

$$= 0,0454545455$$

Kesimpulan : Nilai probabilitas tertinggi yaitu sebesar 0,0454545455 pada  $P(\text{uji}|\text{postif})$  sehingga komentar tersebut diklasifikasikan ke dalam kelas “**Positif**”

3. Menghitung bobot kemunculan *term* terhadap dokumen. Proses pembobotan ini menggunakan rumus *Term Frequency-Invers Document Frequency* (TF-IDF).

**Tabel 4.14** pembobotan data uji

NO	Kata	TF		DF	DF	TF x IDF	
		D1	D2			D1	D2
1	vaksin	1	0	1	0,301029996	0,301029996	
2	aman	1	0	1	0,301029996	0,301029996	
3	bagus	0	1	1	0,301029996		0,301029996

NO	Kata	TF		DF	DF	TF x IDF	
		D1	D2			D1	D2
4	masyarakat	0	1	1	0,301029996		0,301029996
5	hindar	0	1	1	0,301029996		0,301029996
6	covid	0	1	1	0,301029996		0,301029996
7	sebar	0	1	1	0,301029996		0,301029996

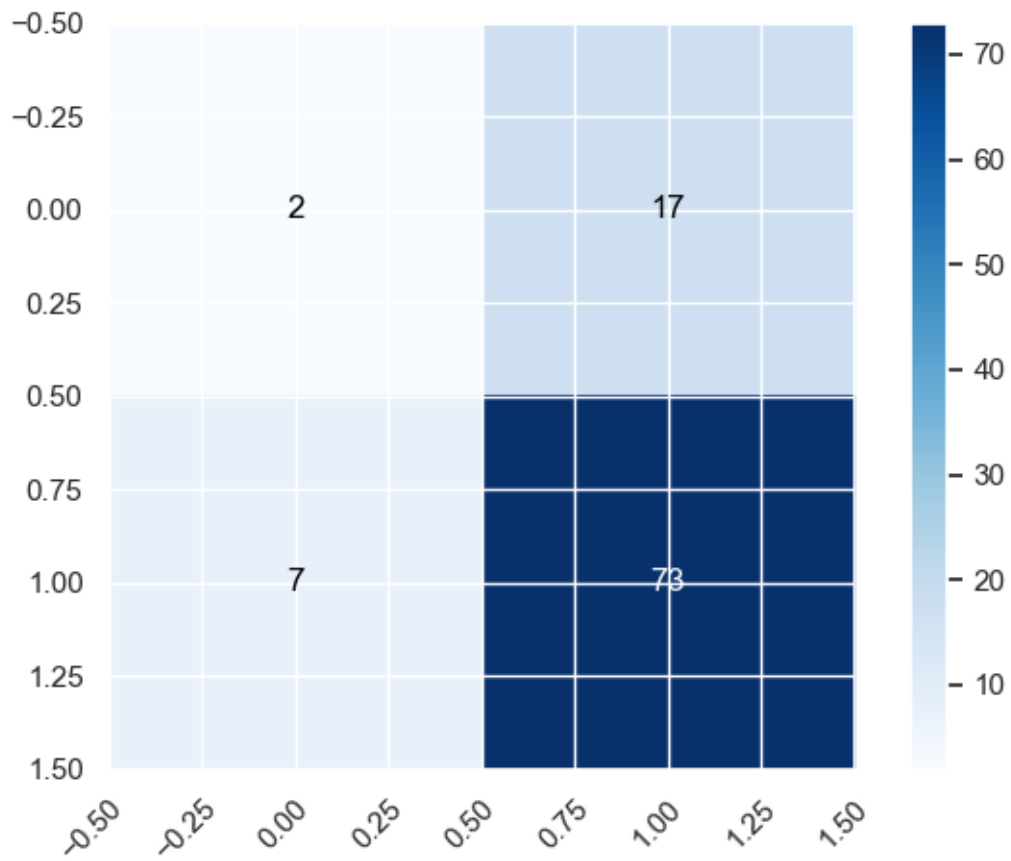
**BAB V**  
**HASIL DAN PEMBAHASAN**

**5.1 HASIL SISTEM**

Setelah proses klasifikasi analisis sentimen selesai dilakukan diperlukan satu langkah untuk menentukan kualitas proses yang telah dilakukan yaitu mengevaluasi hasil dari sistem. Berikut ini merupakan gambar dari Confusion Matrix, without normalization :

[[ 2 17]

[ 7 73]]



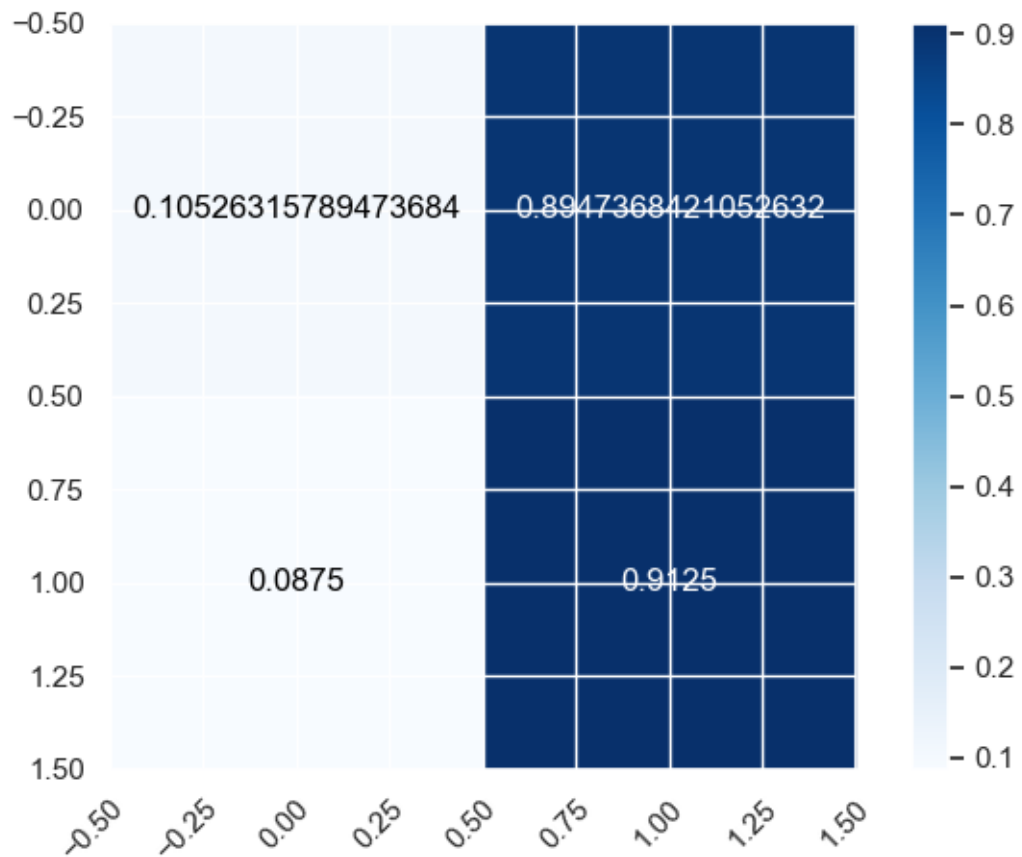
**gambar 5. 1 Confusion Matrix Without Normalization**



Normalized confusion matrix

[[0.11 0.89]

[0.09 0.91]]



**Gambar 5. 2 confusion matrix normalized**

**Tabel 5.1 Hasil confusion Matrix**

Kelas Prediksi	Realita	
	True	False
True	73	7
False	17	2

untuk menguji nilai akurasi dari algoritma naïve bayes maka digunakan rumus pada sub-sub bab 2.2.10 sehingga di dapatkan hasil sebagai berikut:

$$\begin{aligned}
 Accuracy &= \frac{TP + TN}{TP + TN + FP + FN} \\
 &= \frac{73 + 7}{73 + 7 + 17 + 2} \\
 &= 0,8080808081
 \end{aligned}$$

$$\begin{aligned}
 Presicion &= \frac{TP}{TP + FP} \\
 &= \frac{73}{73 + 17} \\
 &= 0,8111111111
 \end{aligned}$$

$$\begin{aligned}
 Recall &= \frac{TP}{FN + TP} \\
 &= \frac{73}{7 + 73} \\
 &= 0,9125
 \end{aligned}$$

[[ 2 17]  
 [ 7 73]]  
 0,9125  
 0,8080808081

gambar 5. 3 Hasil Akurasi



## **BAB VI**

### **KESIMPULAN DAN SARAN**

#### **6.1 Kesimpulan**

1. Berdasarkan hasil klasifikasi menunjukkan bahwa sentimen analisis masyarakat terhadap tindakan vaksinasi covid-19 dominan positif, dimana dari total data 1096, terdapat 616 data termasuk dalam kategori positif, dan 480 data termasuk dalam kategori negative. Sehingga dapat disimpulkan bahwa masyarakat masih menerima tindakan vaksinasi covid-19 dengan baik. Setelah dilakukan klasifikasi dengan menggunakan algoritma Naïve Bayes dan metode TF-IDF akurasi yang dihasilkan adalah baik. Hal ini dapat dilihat dari hasil evaluasi diantaranya adalah tingkat akurasi 80% precision 81% dan recall 91%

#### **6.2 Saran**

Berdasarkan kesimpulan skripsi tersebut diatas, peneliti dapat memberikan saran untuk peneliti selanjutnya;

1. Data yang digunakan adalah data yang memiliki ratio 50:50 data positif dan data negatif sehingga dapat menghasilkan performa sistem yang baik
2. Dapat dikembangkan dengan dikembangkan dengan metode lain, selain metode naïve bayes classifier untuk meningkatkan hasil hasil yang lebih baik

## DAFTAR PUSTAKA

- [1] B. Laurensz and Eko Sedyono, “Analisis Sentimen Masyarakat terhadap Tindakan Vaksinasi dalam Upaya Mengatasi Pandemi Covid-19,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 2, pp. 118–123, 2021, doi: 10.22146/jnteti.v10i2.1421.
- [2] F. F. Rachman and S. Pramana, “Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter,” *Heal. Inf. Manag. J. ISSN*, vol. 8, no. 2, pp. 2655–9129, 2020.
- [3] A. V. Sudiantoro *et al.*, “Analisis Sentimen Twitter Menggunakan Text Mining Dengan Algoritma Naive Bayes Classifier,” *Din. Inform.*, vol. 10, no. 2, pp. 398–401, 2018.
- [4] D. A. Muthia, “Analisis Sentimen Pada Review Restoran Dengan Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes,” *J. Ilmu Pengetah. Dan Teknol. Komput.*, vol. 2, no. 2, pp. 39–45, 2017.
- [5] Sh. Fanissa, M. A. Fauzi, and S. Adinugroho, “Analisis Sentimen Pariwisata di Kota Malang Menggunakan Metode Naive Bayes dan Seleksi Fitur Query Expansion Ranking Optimasi Sisa Bahan Baku Pada Industri Mebel Menggunakan Algoritma Genetika View project Automatic Essay Scoring View project,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 2, no. 8, pp. 2766–2770, 2018, [Online]. Available: <https://www.researchgate.net/publication/322959527>
- [6] Z. L. Argista, *Persepsi Masyarakat Terhadap Vaksin Covid-19 Di Sumatera Selatan: Literature Review*, vol. 13, no. 3. 2021.
- [7] N. Fakhri, “Respon Masyarakat Terhadap Covid-19,” *Pus. Kaji. Psikol. Sos. UNM*, vol. 1, no. April, pp. 1–3, 2020, [Online]. Available: [https://www.researchgate.net/publication/340599055\\_Respon\\_Masyarakat\\_terhadap\\_COVID19](https://www.researchgate.net/publication/340599055_Respon_Masyarakat_terhadap_COVID19)
- [8] I. Rozi, S. Pramono, and E. Dahlan, “Implementasi Opinion Mining (Analisis Sentimen) Untuk Ekstraksi Data Opini Publik Pada Perguruan Tinggi,” *J. EECCIS*, vol. 6, no. 1, pp. 37–43, 2012.

- [9] E. S. Mtsweni *et al.*, “analisis sentimen review pelanggan hotel menggunakan metode k-nearest neighbor,” *Eng. Constr. Archit. Manag.*, vol. 25, no. 1, pp. 1–9, 2020, [Online]. Available: <http://dx.doi.org/10.1016/j.jss.2014.12.010><http://dx.doi.org/10.1016/j.sbspro.2013.03.034><https://www.iiste.org/Journals/index.php/JPID/article/viewFile/19288/19711><http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.678.6911&rep=rep1&type=pdf>
- [10] LAYLA QODARY ZALYHATY, “analisis sentimen tanggapan masyarakat terhadap vaksin covid-19 menggunakan algoritma support vector machine(svm,” p. 6, 2021.
- [11] A. M. Pravina, I. Cholissodin, and P. P. Adikara, “Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme Support Vector Machine ( SVM ),” vol. 3, no. 3, pp. 2789–2797, 2019.
- [12] F. K. Wardani, V. R. Hananto, and V. Nurcahyawati, “Analisis Sentimen Untuk Peningkatan Popularitas Situs Belanja Online Di Indonesia Menggunakan Metode Naive Bayes (Studi Kasus Data Sekunder),” *Jsika*, vol. 08, no. 01, pp. 1–9, 2019.
- [13] N. Fikria, “Analisis Klasifikasi Sentimen Review Aplikasi E-Ticketing Menggunakan Metode Support Vector Machine Dan Asosiasi,” 2018, [Online]. Available: <https://dspace.uui.ac.id/handle/123456789/7717>
- [14] T. Wisata, “1. 439,” pp. 439–448, 2017.
- [15] V. K. S. Que, A. Iriani, and H. D. Purnomo, “Analisis Sentimen Transportasi Online Menggunakan Support Vector Machine Berbasis Particle Swarm Optimization,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 9, no. 2, pp. 162–170, 2020, doi: 10.22146/jnteti.v9i2.102.
- [16] F.Fahriza, "klasifikasi Sentimen Terhadap Gubernur Dki Jakarta Di Media Sosial Twitter Menggunakan Metode Naïve Bayes Classifier,"2021, [Online].Available: <http://repository.uin-suska.ac.id/47385/><http://repository.uin-suska.ac.id/47385/1/SKRIPSI FARID FAHRIZA 11351101165 LENGKAP REVISI.pdf>.



Similarity Report ID: old:25211:32576587

PAPER NAME

**SKRIPSI\_T3118062\_MELISNAWATI I LA  
MARA.docx**

AUTHOR

**Melisnawati Lamara**

WORD COUNT

**5224 Words**

CHARACTER COUNT

**32251 Characters**

PAGE COUNT

**42 Pages**

FILE SIZE

**515.7KB**

SUBMISSION DATE

**Mar 17, 2023 4:57 PM GMT+8**

REPORT DATE

**Mar 17, 2023 4:58 PM GMT+8****● 15% Overall Similarity**

The combined total of all matches, including overlapping sources, for each database.

- 15% Internet database
- 6% Publications database
- Crossref database
- Crossref Posted Content database
- 0% Submitted Works database

**● Excluded from Similarity Report**

- Cited material
- Small Matches (Less than 30 words)

● **15% Overall Similarity**

Top sources found in the following databases:

- 15% Internet database
- 6% Publications database
- Crossref database
- Crossref Posted Content database
- 0% Submitted Works database

TOP SOURCES

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	<b>123dok.com</b> Internet	3%
2	<b>repository.uin-suska.ac.id</b> Internet	2%
3	<b>andi.ddns.net</b> Internet	1%
4	<b>eprints.umm.ac.id</b> Internet	1%
5	<b>id.123dok.com</b> Internet	1%
6	<b>journal.upp.ac.id</b> Internet	1%
7	<b>journal.uinjkt.ac.id</b> Internet	1%
8	<b>journal.unilak.ac.id</b> Internet	1%



9	Siti Lestari, Akmaludin Akmaludin, Mohammad Badrul. "IMPLEMENTA... Crossref	<1%
10	repository.dinamika.ac.id Internet	<1%
11	Santi Thomas, Yuliana, Noviyanti. P. "Study Analisis Metode Analisis ... Crossref	<1%
12	media.neliti.com Internet	<1%
13	sismatik.nusaputra.ac.id Internet	<1%

```
import pandas as pd
import numpy as np
import matplotlib as plt
import seaborn as sns
sns.set(style="darkgrid")
from sklearn.preprocessing import MinMaxScaler
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.feature_extraction.text import TfidfTransformer
from sklearn import feature_extraction, linear_model, model_selection,
    preprocessing
from sklearn.metrics import accuracy_score
from sklearn.model_selection import train_test_split
from sklearn.pipeline import Pipeline
import re
from nltk.stem.porter import PorterStemmer
#from sklearn.feature_extraction.text import TfidfVectorize
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.model_selection import train_test_split
import pickle
from sklearn.linear_model import LogisticRegressionCV
Data = pd.read_excel("Datakotor.xlsx")
Data
Data['komentar'] = Data['komentar'].str.lower()
print('Case Folding Result : \n')
print(Data['komentar'].head(10))
for data in Data:
    data_property = { }
#    data_property['no'] = Data['No']
    data_property['Nama'] = Data['Nama']
    data_property['Alamat'] = Data['Alamat']
    data_property['komentar'] = Data['komentar']
```

```

    print(data_property)
import string
import re #regex Library
# import word_tokenize & FreqDist from NLTK
from nltk.tokenize import word_tokenize
from nltk.probability import FreqDist
def remove_Data(text):
    # remove tab, new line, ans back slice
    text = text.replace('\t'," ").replace('\n'," ").replace('\u'," ").replace('\',"")
    # remove non ASCII (emotion, chinese word, .etc)
    text = text.encode('ascii','replace').decode('ascii')
    # remove mention, link, hashtag
    text = ''.join(re.sub("([@#][A-Za-z0-9+])|(\\w+:\\w+/s+)"," ",text).split())
    # remove incomplete URL
    return text.replace("http://"," ").replace("https://"," ")
Data['komentar'] = Data['komentar'].apply(remove_Data)
len(Data.index)
# remove number
def remove_number(text):
    return re.sub(r"\d+","",text)
Data['komentar'] = Data['komentar'].apply(remove_number)
# remove punctuation
def remove_punctuation(text):
    return text.translate(str.maketrans("", "", string.punctuation))
Data['komentar'] = Data['komentar'].apply(remove_punctuation)
# remove whitespace leading & trailing
def remove_whitespace_LT(text):
    return text.strip()
Data['komentar'] = Data['komentar'].apply(remove_whitespace_LT)

# remove multiple whitespace into singel whitespace

```

```

def remove_whitespace_multiple(text):
    return re.sub("\s+', " ",text)
Data['komentar'] = Data['komentar'].apply(remove_whitespace_multiple)
# remove single char
def remove_single_char(text):
    return re.sub(r"\b[a-zA-Z]\b", "",text)
Data['komentar'] = Data['komentar'].apply(remove_single_char)
# Setup
!pip install -q wordcloud
import wordcloud
import nltk
nltk.download('stopwords')
nltk.download('wordnet')
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')
# NLTK word tokenize
import nltk
nltk.download('punkt')
def word_tokenize_wrapper(text):
    return word_tokenize(text)
Data['komentar'] = Data['komentar'].apply(word_tokenize_wrapper)
print("Tokenizing Result : \n")
print(Data.head(10))
print("\n\n\n")
from nltk.corpus import stopwords
# get stopword from NLTK stopword
# get stopword indonesia
list_stopwords = stopwords.words('indonesian')
# manually add stopword
# append additional stopword
list_stopwords.extend([

```

```

        "yg", "dg", "dgn", "ny", "d", "klo",
        "kalo", "amp", "biar", "bikin", "bilang",
        "gak", "ga", "krn", "nya", "nih", "sih",
        "si", "tau", "tdk", "tuh", "utk", "ya",
        "jd", "jgn", "sdh", "aja", "n", "t",
        "nyg", "hehe", "pen", "u", "nan", "loh", "rt",

        "&", "yah", "jgn", "ga", "ok", "banget", "bgt", "rp", "hrus", "sngtlah", "tdk", "h
        nya", "yg", "sdah", "krna", "ngak"
    ])

# convert list to doctionary
list_stopwords = set(list_stopwords)

# remove stopword pada list token
def remove_stopwords(words):
    return [word for word in words if word not in list_stopwords]

Data['komentar'] = Data['komentar'].apply(remove_stopwords)
print('Stopwords Result : \n')
print(Data['komentar'].head(20))
print("\n\n\n")
pip install Sastrawi
# import Sastrawi package
from Sastrawi.Stemmer.StemmerFactory import StemmerFactory
# create stemmer
factory = StemmerFactory()
stemmer = factory.create_stemmer()
# stemmed
def stemmed_wrapper(term):
    return stemmer.stem(term)

```

```

term_dict = {}
for document in Data['komentar']:
    for term in document:
        if term not in term_dict:
            term_dict[term] = ''
print(len(term_dict))
print("-----")
for term in term_dict:
    term_dict[term] = stemmed_wrapper(term)
    print(term,":",term_dict[term])
print(term_dict)
print("-----")
# apply stemmed term to dataframe
def get_stemmed_term(document):
    return [term_dict[term] for term in document]

Data['komentar'] = Data['komentar'].apply(get_stemmed_term)
print(Data['komentar'].head())
# most frequent words counter (code adapter from
    https://www.kaggle.com/rodolfooluna/fake-news-detector)
from nltk import tokenize
import seaborn as sns
from matplotlib import pyplot as plt
token_space = tokenize.WhitespaceTokenizer()
def counter(saran, columnSaran, quantity):
    all_words = ' '.join([saran for saran in saran[columnSaran].astype('str')])
    token_phrase = token_space.tokenize(all_words)
    frequency = nltk.FreqDist(token_phrase)
    df_frequency = pd.DataFrame({"Word": list(frequency.keys()),
                                "Frequency": list(frequency.values())})
    df_frequency = df_frequency.nlargest(columns = "Frequency", n = quantity)

```

```

plt.figure(figsize=(12,8))
ax = sns.barplot(data = df_frequency, x = "Word", y = "Frequency",
                palette='deep')
# ax = sns.barplot(data = df_frequency, x = "Word", y = "Frequency", color =
                'blue')
ax.set(ylabel = "Count")
plt.xticks(rotation='vertical')
plt.show()
# now many articles per label
print(Data.groupby(['kelas'])['komentar'].count())
Data.groupby(['kelas'])['komentar'].count().plot(kind="bar")
plt.show()
# Data['kelas'].value_counts()
Data['kelas'].value_counts()
from sklearn.linear_model import LogisticRegression
# machine learning naive bayes
from sklearn.naive_bayes import MultinomialNB
from sklearn import model_selection
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.feature_extraction.text import TfidfTransformer
from sklearn.utils.multiclass import unique_labels
from sklearn.model_selection import GridSearchCV
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import scale, LabelBinarizer, OrdinalEncoder,
    OneHotEncoder
# visualisasi hasil
from sklearn.metrics import classification_report, accuracy_score,
    confusion_matrix
import seaborn as sns
import matplotlib.pyplot as plt
xTarget = Data.drop(['kelas','komentar'],axis=1)

```

```

# print(X)
# Xs = scale(X)
yTarget = Data['kelas']
# print(y)
encode = LabelBinarizer()
Y = encode.fit_transform(yTarget)
print(Y)

# X_train, X_test, y_train, y_test = train_test_split(Xs,y,random_state=21,
    stratify=y,test_size=0.2)

# model_lr = LogisticRegression(random_state=21)
# X = DataCLEAN.iloc[:,[1]].values # variabel yang digunakan umur dan salary
    kolom 2 dan 3. semua record (:) kolom 2 dan 3 [2, 3]
# Y = DataCLEAN.iloc[:, -1].values #mengambil data dari semua record (:)
    kolom terakhir (-1)

# x_train, x_test, y_train, y_test = train_test_split(X, Y, test_size=0.25,
    random_state=21)
# model_lr = LogisticRegression(random_state=21)

# param_grid = {
#     'C':np.arange(0,1,1), 'class_weight':[{0:x,1:1-x} for x in
    np.arange(0.1,0.9,0.1)]
# }

# gscv = GridSearchCV(model_lr, param_grid, cv=10,scoring='f1')
# gscv.fit(x_train,y_train)
# merubah (encode) nilai atribut menjadi index nilai
tfidf_transformer = OneHotEncoder()
X = tfidf_transformer.fit_transform(xTarget)

```



```
print(X)
print(X.shape)
x_train, x_test, y_train, y_test = train_test_split(X, Y, test_size=0.30,
            random_state=1)
print(x_train.shape)
print(y_train.shape)
print(x_test.shape)
print(y_test.shape)
# pembuatan model training menggunakan naiveBayes
NaiveBayes = MultinomialNB().fit(x_train,np.ravel(y_train,order='C'))
print(NaiveBayes)
# prediksi terhadap model training
prediction = NaiveBayes.predict(x_test)
accuracies = accuracy_score(y_test, prediction)

print(prediction)
print(accuracies)
# menampilkan matriks confusion
print(confusion_matrix(y_test,prediction))
# function to plot confusion matrix (code from)
# from sklearn import metrics
# from sklearn.metrics import accuracy_score, classification_report,
    confusion_matrix
# import itertools
def plot_confusion_matrix(y_true,y_pred, classes,
            normalize=False,
            title='Confusion Matrix',
            cmap=plt.cm.Blues):
    if not title:
        if normalize:
            title = "Normalized Confusion Matrix"
```

```

    else:
        title = 'Confusion Matrix, without Normalization'
## # compute confusion matrix
    cm = confusion_matrix(y_true,y_pred)
# only use the labels that appear in the data
    classes = classes[unique_labels(y_true,y_pred)]
    if normalize:
        cm = cm.astype('float')/cm.sum(axis=1)[:,np.newaxis]
        print('Normalized confusion matrix')
    else:
        print('Confusion Matrix, without normalization')
    print(cm)
    fig, ax = plt.subplots()
    im = ax.imshow(cm, interpolation='nearest', cmap=cmap)
    ax.figure.colorbar(im,ax=ax)
# we want to show all ticks
# ax.set(xticks=np.arange(cm.shape[1]),
#       yticks=np.arange(cm.shape[0]),
#       xticklabels=classes, yticklabel=classes,
#       title=title,
#       ylabel='True Label',
#       xlabel='Predicted Label') #sementara bermasalah
    plt.setp(ax.get_xticklabels(),rotation=45, ha='right',
             rotation_mode='anchor')
# fmt = '.2f' if normalize else 'd'
    thresh = cm.max()/2.
    for i in range(cm.shape[0]):
        for j in range(cm.shape[1]):
            ax.text(j,i,format(cm[i,j]),
                   ha='center',va='center',
                   color='white' if cm[i,j] > thresh else 'black')

```

```

fig.tight_layout()
return ax
plt.imshow(cm, interpolation='nearest', cmap=cmap)
plt.title(title)
plt.colorbar()
tick_marks = np.arange(len(classes))
plt.xticks(tick_marks, classes, rotation=45)
plt.yticks(tick_marks, classes)
if normalize:
    cm = cm.astype('float') / cm.sum(axis=1)[:, np.newaxis]
    print("Normalized Confusion matrix")
else:
    print('Confusion matrix, without normalization')
print(cm)
thresh = cm.max()/2
for i,j in itertools.product(range(cm.shape[0]), range(cm.shape[1])):
    plt.text(j, i, cm[i,j],
             horizontalalignment="center",
             color="white" if cm[i,j] > thresh else "black")
for i in range(cm.shape[0]):
    for j in range(cm.shape[1]):
        ax.text(j,i,format(cm[i,j],fmt),
               ha='center',va='center',
               color='white' if cm[i,j] > thresh else 'black')
plt.tight_layout()
plt.ylabel('True label')
plt.xlabel('Predicted label')
class_names = yTarget
np.set_printoptions(precision=2)
# plot Non-Normalize confusion matrix
plot_confusion_matrix(y_test, prediction, classes=class_names,

```

```
        title='Confusion matrix, without Normalization')  
plt.show()  
# plot Normalize confusion matrix  
plot_confusion_matrix(y_test, prediction, classes=class_names,normalize=True,  
        title='Normalize confusion matrix')  
  
plt.show()
```

## Dataset

No	Nama	Alamat	Komentar
1	isti irmawati hanapi	desa liodu	bicara covid menurut sekarang bincang juta umat wkwkwkwkbahkan kenal bagi virus berbahaya makanya adakan vaksin apapun manfaat vaksin tubuh suntik vaksin rangsan antibodi ajar mengenali virus lemah tubuh virus kurang risiko terpaparnah saya tujuan adanya vaksin sekarang vaksin tahap mas rentan terkenah covid yang kenal virus omicron nah lantas guna vaksin tubuh mas papar virus covid
2	Wilen	Desa liodu	resah masyarakat urus data data vaksinasi
3	Yeyen Hemeto	Desa Liyodu	tindak vaksinasi efektif vaksin daya tahan tubuh rentan orang vaksinasi kena covid lantas laku vaksin vaksin vaksin kena covid
4	Rolan ma'ruf	Desa liyodu kec. Bongomem e	vaksinhalal

5	Maryam Lasena	Desa liyodu	masyarakat dapat paksa vaksinasi
6	Saida K. Dukala ng	Desa Liyodu, Kec.Bongo meme	setuju vaksin
7	Sutika	Desa Liyodu	dapat vaksinasi tindak coba tubuh manusia bakteri dalam tubuh pengaruh bugar tubuh tubuh manusia rentan bakteri dalam tubuh manusia kurang tuju vaksinasi vaksinasi administrasi
8	Hadijah lasena	Desa liyodu	lindung papar covid
9	Hatu k.lasena	Desa liyodu	ragu vaksin karna efek samping
10	Kuyu lasena	Desa liyodu	vaksin aman halal
11	Ngoi maini	Desa liyodu	tidak setuju vaksin karena takut efek samping
12	Indrawati pomu	Desa Liyodu Kec Bongomem e	resah vaksin sakit sakitan sesak nafas
13	Wisnawati s polioto	Desa liyodu	bagus bantu cegah tular virus covid
14	Ismail lamara	Liyodu, bongomeme Gorontalo	baik untuk masyarakat
15	Rusmin poiyo	Desa liyodu	vaksin aman halal

16	Yunus Lasena	Desa liyodu	syukur vaksin karena vaksin jaga tubuh virus
17	ummul hasanah	desa dulamayo	tangkal covid
18	Santi	Desa liyodu	bekeng ribet
19	ummul	Desa liodu	vaksin meribetkan mahasiswa pulang kampung
20	annisa ramadh any s.	desa liyodu	masyarakat vaksin dosis perintah harus masyarakat ikut tindak vaksin dosis tiga
21	Wulan angii	Desa huntulohula wa	dapat saya tindak vaksinasi covid
22	Ningsi lasena	Desa liyodu kec.bongom eme kab.goronta lo	dunia alami krisis sehat ekonomi
23	Nangsi lasena	Desa liyodu kec.bongom eme kab gorontalo	dunia alami krisis sehat ekonomi
24	Fatma poiyo	Desa liyodu	vaksin aman vaksin halal
25	Mustapa maini	Desa liyodu	vaksin aman halal
26	Sandra ngguyu	Desa liyodu	vaksin aman
27	Ical lasena	Desa liyodu	alhamdulillah vaksin aman halal

28	Jufri lasena	Desa liyodu	saya tidak percaya vaksin
29	Nurhadijah Luawo	Desa liyodu	vaksin aman
30	Desi	Desa liyodu	saya bagus vaksin tambah kebal tubuh
31	Desriwanti Djauhari	Desa Liyodu kec bongomeme	tindak covid jaga sehat jaga selamat dampak virus covid mematahui protokol sehat jaga jarak pakai masker cuci tangan makan tindak covid vaksin vaksin masuk vaksin vaksin jalan pasien covid kurang pasien covid tambah jaga sehat
32	Nurnaningsih I Kadoli	Desa liyodu	dapat tindak vaksin covid mau paksa desa
33	Melisa Lamara	Liyodu	vaksin tidak penting
34	Santi Harun	Desa Liyodu	vaksinasi covid paksa
35	RIFKA	Desa Liyodu	dapat tindak vaksinasi covid unsur paksa vaksin ancam salah satu sulit urus suratsurat
36	Lartin	Liyodu	tindak vaksinasi covid efektif
37	Sartin	Liyodu	pandang vaksinasi covid perintah dinas sehat suruh vaksin ikut atur



			rakyat sulit
38	Warni	Liyodu	tindak vaksinasi covid sehat
39	Intan	Liyodu	vaksinasi covid bagus karena vaksin haram
40	Heni	Liyodu	vaksinasi covid pengaruh tubuh
41	Onal	Desa liyodu	bekeng saki badan
42	Indah	Liyodu	vaksin efektif
43	Desriyanti lobo	Desa liyodu	paham fungsi manfaat vaksin
44	Chikq deriska sembag a	Desa liyodu	vaksin baik
45	Indi	Liyodu	tuju vaksin
46	Dea Badu	Desa Liyodu	vaksin lindung papar covid
47	NIKSON DJAFAR	Desa liyodu kec bongomeme	bantu kurang angka mati karena covid
48	Noho upohula wa	Desa liyodu	vaksin aman
49	Hasna wadiko	Desa liyodu	vaksin aman
50	Elan hubulo	Desa liyodu	vaksin aman halal
51	Adinda Putri Dehi	Desa liyodu	badan sehat lindung virus bahaya
52	Elen hubulo	Desa liyodu	vaksin aman
53	Yuyun lamara	Desa liyodu	alhamdulillah sehat sejatra
54	Mohammad	Desa liyodu	vaksin aman

	hubulo		
55	Yuyun lamara	Desa liyodu	alhamdulillah sehat sejahtera
56	ANISA MUSA	Kab Gorontalo	bagus
57	Pipin	Desa liyodu	bentuk anti body
58	Lilis	Liyodu	dapat tindak vaksin covid vaksin laku indonesia pentingdi karena salah putus rantai kovid indonesia meskipun skarang salah orang kovid dan jugah percaya vaksin anjur perintah daerahtapi bejalannya waktuh masyarakat sadar kovid menceganya denga vaksin nasi
59	Afrilyanti Harun	Desa Liyodu	tindak vaksinasi lindung papar virus covid
60	Inul Hemeto	Desa Liyodu	tindak vaksinasi covid langkah baik ambil perintah upaya perintah indonesia tangan covid vaksinasi covid tuju cipta kebal kelompok herd immunity masyarakat produktif jalan aktivitas hari
61	Risti fadilah	Desa liyodu	vaksinasi dapat cegah kena

			covid
62	Egy Lasena	Desa Liyodu	vaksinasi jaga daya tahan tubuh kuat
63	Dewy Lamara	Desa Liyodu, kec Bongomeme	vaksin efektif
64	Nurwigya Mohamad Putri	Desa Liyodu	vaksinasi covid tindak meminimalisir tular covid
65	Nurmelinda A. Popa	Desa Liyodu	vaksinasi sasar cepat
66	Nurfauzia N. Luawo	Desa Liyodu	dengar tri sehat vaksinasi covid bantu tahan kebal tubuh sistem imun manusia
67	Zunaidi K. Gagulu	Desa Liyodu	baik
68	Harmianti kasno	Liyodu	rugi
69	Yusnita Antu	Desa Liyodu	vaksinasi pandemi covid tekan tambah vaksinasi tuju sistem kebal tubuh nali cepat lawan bakteri virus sebab infeksi tuju capai beri vaksin covid turun angka sakit angka mati akibat virus
70	Mekka Aleesya Ismail	Desa Liyodu	bermanfaat

71	Putra Arif Yanto Umar	desa dulamayo	lumayan tingkat vaksinasi gratis moga depan covid omicron virus julang vaksinasi jalan lancar
72	Febrianti A. Husain	Desa liyodu	vaksin sangat efektif
73	Tinsy	Desa liyodu	vaksin bermanfaat
74	Nay hemeto	Desa Liyodu	moga covid cepat
75	Reski Rahmat ia	Desa Liyodu	vaksinasi efektif
76	Rizky Pauwen i	Desa Liyodu	bantu vaksin aman virus covid
77	Elvin paulu	Desa liyodu	tindak vaksinasi lindung papar covid namun raguragu setuju mereka nilai vaksin covid miliki efek samping
78	Rastra pratiwy tolinggi	Desa liyodu	bagus masyarakat hindar covid sebar ini
79	Maryam	Desa liyodu	efektif
80	Nai	Desa liyodu	vaksin baik
81	Yuyun gusasi	Desa liyodu	vaksin repot
82	Fira	Desa liyodu	vaksin untuk kebal tubuh
83	Sella poiyo	Desa liyodu	vaksin efektif
84	Anto abas	Desa liyodu	vaksin aman badan
85	Murni	Desa liyodu	vaksin aman halal
86	Linda	Desa liyodu	vaksin bisnis saja

87	Agus	Desa liyodu	vaksin aman badan
88	Fatma matahul i	Desa liyodu	vaksin tidak baik
89	Yayu	Desa liyodu	vaksin badan
90	Rui	Desa liyodu	vaksin langsung rasa saki demam
91	Aldo akuba	Desa liyodu	vaksin aman
92	Lilan	Desa liyodu	vaksin baik
93	Mila	Desa liyodu	vaksin langsung sakit
94	Nina	Desa liyodu	vaksin aman
95	Maryam hula'o	Desa liyodu	mas takut vaksin
96	Sari	Desa liyodu	vaksin tidak aman
97	Epi	Desa liyodu	vaksin efektif
98	Nuni	Desa liyodu	vaksin aman halal
99	Nangsi lasena	Desa liyodu	vaksin rugi
100	Tiara hemeto	Desa liyodu	vaksin
.....	.....	.....	.....
1096	sri fani G lamara	desa molanihu	vaksin covid wajib karna mencegah sakit



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI  
**UNIVERSITAS ICHSAN GORONTALO**  
**FAKULTAS ILMU KOMPUTER**  
**SURAT KEPUTUSAN MENDIKNAS RI NOMOR 84/D/O/2001**  
 Jl. Achmad Najamuddin No. 17 Telp. (0435) 829975 Fax (0435) 829976 Gorontalo

**SURAT REKOMENDASI BEBAS PLAGIASI**  
 No. 108 /FIKOM-UIG/R/III/2023

Yang bertanda tangan di bawah ini :

Nama : Irvan Abraham Salihi, M.Kom  
 NIDN : 0928028101  
 Jabatan : Dekan Fakultas Ilmu Komputer

Dengan ini menerangkan bahwa :

Nama Mahasiswa : Melisnawati I. Lamara  
 NIM : T3118062  
 Program Studi : Teknik Informatika (S1)  
 Fakultas : Fakultas Ilmu Komputer  
 Judul Skripsi : Analisis Sentimen Masyarakat Terhadap Tindakan  
 Vaksinasi Covid 19 Menggunakan Metode Naive  
 Bayes

Sesuai hasil pengecekan tingkat kemiripan skripsi melalui aplikasi **Turnitin** untuk judul skripsi di atas diperoleh hasil *Similarity* sebesar **15%**, berdasarkan Peraturan Rektor No. 32 Tahun 2019 tentang Pendeteksian Plagiat pada Setiap Karya Ilmiah di Lingkungan Universitas Ichsan Gorontalo dan persyaratan pemberian surat rekomendasi verifikasi calon wisudawan dari LLDIKTI Wil. XVI, bahwa batas kemiripan skripsi maksimal 30%, untuk itu skripsi tersebut di atas dinyatakan **BEBAS PLAGIASI** dan layak untuk diujikan.

Demikian surat rekomendasi ini dibuat untuk digunakan sebagaimana mestinya.

Mengetahui  
 Dekan,  
  
**Irvan Abraham Salihi, M.Kom**  
 NIDN. 0928028101

Gorontalo, 27 Maret 2023  
 Tim Verifikasi,

  
**Zulfrianto V. Lamasigi, M.Kom**  
 NIDN. 0914089101

Terlampir :  
 Hasil Pengecekan Turnitin



KEMENTERIAN PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI  
**UNIVERSITAS ICHSAN GORONTALO**  
FAKULTAS ILMU KOMPUTER  
UPT. PERPUSTAKAAN FAKULTAS  
**SK. MENDIKNAS RI NO. 84/D/0/2001**  
**Jl. Achmad Nadjamuddin No.17 Telp(0435) 829975 Fax. (0435) 829976 Gorontalo**

**SURAT KETERANGAN BEBAS PUSTAKA**

No : 003/Perpustakaan-Fikom/XII/2022

Perpustakaan Fakultas Ilmu Komputer (FIKOM) Universitas Ichsan Gorontalo dengan ini menerangkan bahwa :

Nama Anggota : Melisnawati I. Lamara  
No. Induk : T3118062  
No. Anggota : M2022130

Terhitung mulai hari, tanggal : Sabtu, 10 Desember 2022, dinyatakan telah bebas pinjam buku dan koleksi perpustakaan lainnya.

Demikian keterangan ini di buat untuk di pergunakan sebagaimana mestinya.

**Gorontalo, 10 Desember 2022**

**Mengetahui,  
Kepala Perpustakaan**



**Apriyanto Alhamad, M.Kom**  
**NIDN : 0924048601**



**PEMERINTAH KABUPATEN GORONTALO  
KECAMATAN BONGOMEME**

*Jl. Nani Hasan-Dulamayo Kodepos 96271*

**REKOMENDASI**

Nomor : 074/K.BGM/554/XII/2022

Yang bertanda tangan dibawah ini :

N a m a : ASMIN POPA, S.IP, M.Si  
NIP : 19790517 199903 2 001  
Jabatan : PLT. CAMAT BONGOMEME

Menerangkan bahwa :

N a m a : MELISNAWATI I. LAMARA  
NIM : T3118062  
Fakultas : Ilmu Komputer  
Program Studi : Teknik Informatika  
Judul Penelitian : Analisis Sentimen Pro dan Kontra Terhadap Tindakan Vaksinasi Covid-19

Bahwa yang Bersangkutan benar-benar telah selesai melakukan pengambilan data penelitian dalam rangka Penyusunan Proposal/skripsi di Kantor Camat Bongomeme Kabupaten Gorontalo.

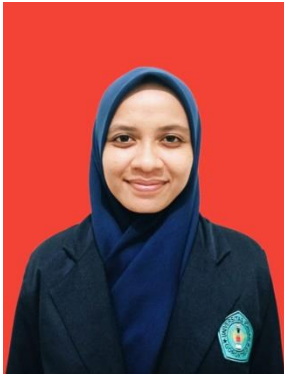
Demikian Surat keterangan ini di buat dengan sebenarnya untuk dapat di gunakan seperlunya.

Dikeluarkan di : Bongomeme  
Pada Tanggal : 08 Desember 2022

PLT. CAMAT BONGOMEME   
  
**ASMIN POPA, S.IP, M.Si**  
NIP. 19790517 199903 2 001



## Daftar Riwayat Hidup



**Nama** : Melisnawati I Lamara  
**Nim** : T3118062  
**Tempat, Tanggal Lahir** : Bongomeme, 18 Mei 1998  
**Alamat** : Jln. Nani Hasan  
**Agama** : Islam  
**Kewarganegaraan** : WNI  
**Email** : lissalamara643@gmail.com

### Riwayat pendidikan dan pekerjaan :

1. Tahun 2011, menyelesaikan Pendidikan di Sekolah Dasar Negeri 17 Wonosari.
2. Tahun 2014, menyelesaikan Pendidikan di Sekolah Menengah Pertama Negeri 1 Bongomeme Kabupaten Gorontalo.
3. Tahun 2017, menyelesaikan Pendidikan di Sekolah Menengah Kejuruan Negeri 1 Batudaa.
4. Tahun 2018, diterima menjadi mahasiswa di Perguruan Tinggi Swasta Universitas Ichsan Gorontalo.

