

# 3376-8988-1-SM

*by jiko unkhair*

---

**Submission date:** 14-Aug-2021 09:38AM (UTC+0200)

**Submission ID:** 1626265437

**File name:** 3376-8988-1-SM.docx (165.19K)

**Word count:** 2852

**Character count:** 18432

## ANALISIS SENTIMEN PENGGUNA YOUTUBE TERHADAP TAYANGAN #MATANAJWAMENANTITERAWAN DENGAN METODE NAÏVE BAYES CLASSIFIER

Mathoril Hudha<sup>1</sup>, Endang Supriyati<sup>2</sup>, Tri Listyorini<sup>3</sup>

<sup>1,2,3</sup>Teknik Informatika, Fakultas Teknik, Universitas Muria Kudus  
Email: <sup>1</sup>mathoril.hudha@gmail.com, <sup>2</sup>endang.supriyati@umk.ac.id, <sup>3</sup>trilistyorini@umk.ac.id

(Naskah masuk: 04 Agustus 2021, diterima untuk diterbitkan: 01 Oktober 2021)

### Abstrak

Tayangan #MataNajwaMenantiTerawan menuai pro dan kontra setelah Najwa Shihab dalam acara Mata Najwa mewawancarai kursi kosong yang direpresentasikan sebagai Menteri Kesehatan Terawan Agus Putranto. Wawancara dengan kursi kosong biasa dilakukan di luar negeri yang memiliki histori pers yang lama. Seperti yang dilakukan Piers Morgan di CNN dan Kay Burley di acara Kay Burley Show di Sky News. Media sosial mengalami peningkatan penggunaan sehingga analisis senti perlu dilakukan untuk mengetahui opini publik terhadap isu yang sedang terjadi. Analisis sentimen dalam penelitian ini dilakukan untuk mengetahui bagaimana penilaian publik terhadap tayangan #MataNajwaMenantiTerawan di Youtube. Data dalam penelitian ini diambil dengan cara scrapping dari komentar video youtube #MataNajwaMenantiTerawan. Dataset mentah hasil scrapping dilakukan preprocessing untuk memudahkan proses klasifikasi. Algoritma Naive Bayes digunakan untuk mengklasifikasikan komentar ke kategori sentimen positif, sentimen negatif dan netral. Algoritma Naive Bayes menghasilkan akurasi sebesar 90.36%. Sentimen netral lebih mendominasi dengan jumlah 1232 data netral, 90 data negatif dan 78 data positif.

**Kata kunci:** Analisis Sentimen, Naive Bayes, Text Preprocessing, Scrapping, #MataNajwaMenantiTerawan

## SENTIMENT ANALYSIS OF USER COMMENTS ON YOUTUBE VIDEO #MATANAJWAMENANTITERAWAN WITH NAIVE BAYES CLASSIFIER METHOD

### Abstract

The show #MataNajwaMenantiTerawan reaped pros and cons after Najwa Shihab in the Mata Najwa program interviewed an empty chair represented as Minister of Health Terawan Agus Putranto. Interviews with empty chair are usually conducted abroad, which have a long press history. Like Piers Morgan on CNN and Kay Burley on the Kay Burley Show on Sky News. Social media has increased its use so that sentiment analysis needs to be done to find out public opinion on the issues that are happening. Sentiment analysis in this study was conducted to find out how the public evaluates the #MataNajwaMenantiTerawan show on Youtube. The data in this study were taken by scrapping from comments on the youtube video #MataNajwaMenantiTerawan. The raw dataset from scrapping is preprocessed to facilitate the classification process. Naive Bayes algorithm is used to classify comments into categories of positive sentiment, negative sentiment and neutral. The Naive Bayes algorithm produces an accuracy of 90.36%. Neutral sentiment dominates with 1232 neutral data, 90 negative data and 78 positive data.

**Keywords:** Analisis Sentimen, Naive Bayes, Text Preprocessing, Scrapping, #MataNajwaMenantiTerawan

### 1. PENDAHULUAN

Pers merupakan subsistem dari sistem sosial. Bagi bangsa Barat perkembangan pers dan teknologinya mengikuti perkembangan masyarakatnya yang berjalan secara linear. Ketika mesin cetak ditemukan pada 1450 sampai ditemukan lagi kamera/film 400 tahun kemudian, tumbuhlah budaya membaca dan Pendidikan yang meluas dan mapan. 50 tahun setelah teknologi visual, pers masuk ke media audio dengan ditemukannya pemancar

radio. Masa-masa berikutnya teknologi pers dan penyiaran terus diperpendek dengan ditemukannya media berbasis audio-visual yang dikenal dengan tabung gambar atau televisi 30 tahun berikutnya. Disusul 10 tahun berikutnya masyarakat barat sampai kepada teknologi satelit komunikasi[1].

Di era *Internet of Things* (IoT) ini perkembangan teknologi tersebut dapat kita nikmati melalui satu perangkat smartphone. Apple, Windows dan Android menyediakan berbagai platform untuk

menikmati konten-konten digital. Kita hanya perlu menginstall aplikasi dan menyediakan kuota data internet untuk mengakses konten yang dimaksud. Seperti jika kita ingin menikmati konten visual seperti berita teks atau info grafis melalui *web browser*, konten audio melalui *Spotify* atau konten audio-visual melalui *Youtube*.

*Nielsen Cross-Platform* pada tahun 2017 melakukan survei terhadap pengguna internet. Survei tersebut menunjukkan terjadinya peningkatan akses internet oleh netizen di banyak tempat. Di antaranya adalah (51%) di kafe atau restoran dan (53%) kendaraan umum. Acara konser pun mengalami peningkatan (24%) dalam jumlah akses media digital dibanding tahun 2015. Rumah dan tempat kerja juga mengalami peningkatan akses internet. Berbagai portal berita seperti *jawapos.com*, *detik.com*, *kompas.com*, *tribun.com* dan *tempo.co* merupakan contoh media yang menyesuaikan dengan berkembangnya teknologi. Semula media tersebut hanya menyajikan produk cetak berupa koran, majalah atau buku. Tayangan televisi pun menyesuaikan perkembangan teknologi seiring masuknya internet. Salah satu layanan penyiaran populer berbasis audio-visual adalah *Youtube*. Pengguna *Youtube* mencapai 95% dari pengguna internet. Konten yang tersedia di *Youtube* memiliki 80 bahasa yang berbeda. Lebih dari 2 miliar pengguna baru mengunjungi *Youtube* dalam setiap bulan dan videonya ditonton lebih dari satu miliar jam setiap hari[2]. Selain sebagai media penyiaran, *Youtube* juga media sosial yang memungkinkan komunikasi antar penggunanya. *Youtube* adalah salah satu platform yang banyak dimanfaatkan TV untuk memublish siarannya. Seperti Kick Andy di MetroTV, Indonesia Lawyers club di TvOne, dan Aiman di Kompas TV yang juga memublish tayangannya di *Youtube*.

Di masa pandemi seperti saat ini, masyarakat memerlukan informasi dari sumber terpercaya untuk menghadapi pandemi. Najwa Shihab melalui program Mata Najwa di Trans TV berusaha mencari kejelasan informasi dari Menteri Kesehatan Terawan Agus Putranto. Wawancara tersebut juga di publish melalui kanal *Youtube* Najwa Shihab setelah acara tersebut ditayangkan di TV secara live. Dalam acara tersebut Najwa Shihab mewawancarai kursi kosong yang direpresentasikan sebagai Menteri Kesehatan saat Pandemi COVID-19 mewabah. Wawancara tersebut dilakukan Najwa Shihab karena Menteri Kesehatan sangat jarang memberikan keterangan mengenai pandemi COVID-19.

Wawancara dengan kursi kosong biasa dilakukan di luar negeri yang memiliki histori pers yang lama. Seperti yang dilakukan Piers Morgan di CNN dan Kay Burley di acara Kay Burley Show di Sky News. Namun Hal tersebut belum pernah dilakukan di Indonesia sehingga menuai pro dan kontra yang memicu penonton tayangan #MataNajwaMenantiTerawan mengungkapkan

emosinya. Emosi yang diluapkan di kolom komentar tayangan #MataNajwaMenantiTerawan inilah yang menjadi fokus peneliti melakukan analisis sentimen dengan metode *Naïve Bayes Classifier*.

## 2. Tinjauan Pustaka

### 2.1. Penelitian Terdahulu

Penelitian mengenai analisis sentimen telah dilakukan oleh Ghulam Asrofi Buntoro dengan studi kasus Tweet mengenai Calon Gubernur DKI Jakarta. Pada penelitian tersebut Ghulam melakukan klasifikasi sentimen dengan menggunakan Metode *Naïve Bayes Classifier*. Penelitian ini menggunakan metode *Naïve Bayes Classifier* dalam Bahasa Indonesia dengan kata kunci AHY, Ahok dan Anies dengan dataset sebanyak 300 tweet. Hasil dari klasifikasi dengan *Naïve Bayes Classifier* diperoleh akurasi lebih tinggi dari *Support Vector Machine (SVM)* dengan dengan nilai rata-rata akurasi 95%. [3].

Penelitian yang dilakukan oleh Christianto dkk, memaparkan proses yang dilakukan dalam analisis sentiment komentar berbahasa Indonesia di *Youtube* merupakan proses mengolah data mentah yang telah dikumpulkan dari *Youtube*. Sebelum dilakukan training dan klasifikasi dengan metode *Naïve Bayes Classifier*. *Preprocessing* perlu dilakukan untuk mengatasi data mentah yang bervariasi dan tidak konsisten sehingga dapat mempengaruhi proses training dan proses klasifikasi. Proses klasifikasi adalah proses dimana algoritma klasifikasi dijalankan kepada komentar yang digunakan sebagai input data komentar. Hasil dari penelitian tersebut menghasilkan nilai *smoothness*. Nilai tersebut menunjukkan bahwa komentar termasuk ke dalam pengelompokan *class* sentimen positif, sentimen negatif atau netral[4].

Saputra melakukan *opinion mining* terhadap video pelayanan pemerintah di *Youtube*. Komentar pada video merupakan data yang nantinya diklasifikasikan menjadi klasifikasi positif, negatif dan netral. Diperoleh nilai akurasi sebesar 69.23% untuk *opinion mining* terhadap KemenPUPR dan 64.10% untuk *opinion mining* terhadap Kemenkeu[5].

Penelitian oleh Nugroho menggunakan 800 data tweet dari pengguna jasa ojek online dengan 300 *data training* dan 500 *data testing*. Penelitian tersebut menghasilkan sebuah sistem *opinion mining* terhadap pengguna jasa ojek online dengan metode *Naïve Bayes*[6].

Samsir pada awal November 2020 melakukan *sentiment analysis* terhadap pembelajaran tatap maya saat pandemi COVID-19 di Indonesia. Data dalam penelitian tersebut diambil dengan cara *text mining* pada *Twitter*. Data yang telah melalui *preprocessing* dianalisis menggunakan algoritma *Naïve Bayes*. Hasil klasifikasi menunjukkan pembelajaran daring memiliki 30% sentimen positif, 69% sentimen negatif dan 1% netral. Beberapa tweet

menunjukkan ketidakpuasan pembelajaran daring. Kata-kata seperti 'stres' dan 'malas' banyak masuk ke dalam frekuensi kata yang sering muncul yang diklasifikasikan sebagai sentimen negatif[7].

Puji dkk, melakukan analisis sentimen terhadap penggemar K-POP dengan algoritma *Naïve Bayes*. Hasil analisis menunjukkan 58,5% sentimen netral, 7,3% sentimen positif dan 34,2% sentimen negatif dari 1000 data tweet. Akurasi algoritma *Naïve Bayes* dalam penelitian tersebut mencapai 75%[8].

Penerapan *data science* dilakukan oleh Balya dalam analisis sentimen pengguna *Youtube* terhadap video review *smartphone*. Salah satu penerapannya adalah *Natural Language Processing* dan *Text Mining* atau dikenal juga dengan istilah *text analytics*. Metode dalam penelitian tersebut menggunakan algoritma *Naive Bayes*. Proses dari penelitian tersebut adalah dengan melakukan *text mining* pada situs *Youtube* mengenai review *smartphone* dengan *scrapping*, labelisasi dan *preprocessing* (*case folding, tokenizing, filtering*), perhitungan frekuensi kemunculan kata atau *Term Frequency (TF)* dan klasifikasi sentimen. Hasil dari penelitian tersebut adalah pengklasifikasian komentar pada video *Youtube* dalam sentimen positif, negatif dan netral[9].

## 2.2. Analisis Sentimen

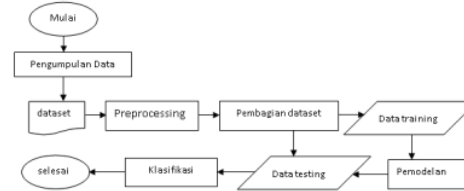
Untuk menemukan opini dari penulis tentang suatu entitas tertentu dilakukan dengan analisis sentimen. Analisis sentimen di media social lebih sulit dengan jenis teks lain karena banyaknya ketidak sempurnaan pada teks seperti singkatan dan jargon. Namun demikian, media sosial memberikan lebih banyak informasi di luar teks seperti media terkait, reaksi pengguna dan hubungan antar pengguna [10].

*Opinion mining* atau analisis sentimen merujuk pada bidang *text mining*, komputasi linguistic dan pengolahan bahasa alami. Dalam arti harfiah, *opinion mining* bertujuan untuk menilai *attitude* dari pembuat opini yang berhubungan dengan suatu topik[11]. Analisis yang dilakukan fokus pada sentimen positif, negatif dan netral dengan pengklasifikasi ketiga kelas sentimen dengan algoritma analisis sentimen[12].

## 3. METODE

### 3.1. Diagram Alur Proses

Penelitian dilakukan dengan mengumpulkan komentar *Youtube* tayangan #MataNajwaMenantiTerawan. Komentar tersebut merupakan data yang akan dianalisis dalam penelitian ini. Berikut ini adalah jalannya penelitian yang digambarkan di Gambar 1.



Gambar 1. Diagram Alur Proses

Gambar 1 menjelaskan tentang alur secara umum yang dilakukan dalam penelitian ini. Penelitian ini dimulai dengan pengumpulan data yang berupa teks komentar. Data tersebut dilakukan preprocessing, vektorisasi, pembagian data, pemodelan dan klasifikasi.

### 3.2. Pengumpulan Data

Penelitian ini mengambil data dengan cara *scrapping* komentar di *Youtube* tayangan #MataNajwaMenantiTerawan. *Scrapping* dilakukan dengan memanfaatkan *library Selenium* dengan bahasa pemrograman Python. *Chrome driver* digunakan untuk membuka google chrome dan mengendalikannya secara otomatis. Data yang telah diambil dari proses *scrapping* disimpan dalam bentuk CSV yang nantinya akan diproses lebih lanjut.

### 3.3. Preprocessing

Data yang tidak terstruktur dan tidak konsisten perlu dibersihkan melalui proses *text preprocessing*. Setelah data bersih maka akan dilakukan proses pelabelan secara manual[13]. Pada proses ini digunakan *googlesheet* dengan *add-ons Logic sheet* untuk melakukan *Cleansing, Case Folding, Tokenize, Stopword Removal* dan *Stemming*. Berikut ini adalah penjelasan dari tahap *preprocessing*:

**Cleansing:** Dataset mentah dibersihkan dari symbol, emoji, tanda baca, link url, tanda baca dan angka.

**Case Folding:** Dataset yang memuat huruf besar diubah menjadi huruf kecil.

**Tokenize:** Kalimat komentar dipecah menjadi kata per kata seperti "dia sedang mewawancarai kursi" menjadi "dia", "sedang", "mewawancarai", "kursi".

**Stopword Removal:** Kata hubung atau kata-kata yang tidak penting dalam kalimat, misalnya "di", "ke", "dari" atau "yang".

**Stemming:** Pada proses ini dilakukan perubahan kata-kata yang memiliki imbuhan seperti "mainan" menjadi "main".

### 3.4. Naïve Bayes Classifier

Metode yang simpel dengan hasil yang baik untuk mengklasifikasi teks adalah keunggulan *Naïve Bayes Classifier*. Keunggulan lain melakukan klasifikasi dengan *Naïve Bayes* adalah cepat dan akurasi yang tinggi [18, 14]. Berikut ini adalah rumus *Naïve Bayes* dalam persamaan (1).

$$P(A|S) = \frac{P(S|A) \cdot P(A)}{P(S)} \quad (1)$$

Keterangan:

- P(A|S) : Probabilitas A terhadap kondisi S
- P(S|A) : Probabilitas S terhadap kondisi A
- P(A) : Probabilitas A
- P(S) : Probabilitas S

## 4. HASIL DAN PEMBAHASAN

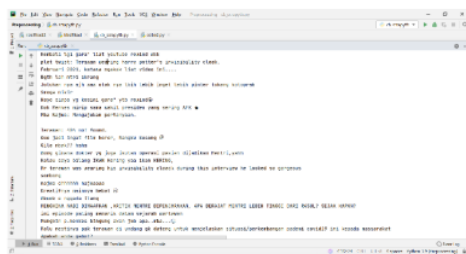
### 4.1. Scrapping

Data dalam penelitian ini diperoleh dari komentar pengguna *Youtube* dalam tayangan #MataNajwaMenantiTerawan. Data tersebut diambil dengan cara scrapping menggunakan Bahasa pemrograman python dengan memanfaatkan *library selenium, chrome driver* dan *xpath*. Dalam penelitian ini diambil data sebanyak 1400 komentar dalam waktu 1 menit. Komentar diambil pada tanggal 8 Juli 2021. Hasil *scrapping* dipaparkan di Tabel 1.

Tabel 1. Hasil *Scrapping*

Komentar
krispi banget, ngomong sama mahluk ghoib
Yang kesini gara gara YouTube rewind like 🤔🤔
Menteri taunya cuma duit doang n korupsi Prestasi nihil
BPJS doesn't cover that burn, bro.
Semoga Indonesia lebih baik dengan adanya orang orang seperti mba nana untuk kritik yang membangun □ pertanyaan bagus loh, jdi penasaran bgt jawabannya, sayang gabisa hadir
Hahahaha... 🤔🤔🤔
Gile yang komen 58 ribu.. luar biasa mba nana
Plot twist : Terawan is just the chair's name
pak?? kemana pak meninggoy

Hasil *Scrapping* juga dapat dilihat dari jalannya program di Gambar 2 sebagai berikut:



Gambar 2. Hasil *Scrapping*

### 4.2. Text Preprocessing

Sebelum data diproses dengan Algoritma *Naïve Bayes*, data harus melewati tahap *preprocessing*. Data yang tidak konsisten perlu dilakukan *preprocessing* agar menjadikan data siap untuk diproses[5]. *Preprocessing* terdiri dari langkah *Tokenizing, Normalisasi, Stemming* dan Menghapus *Stopword*. Pada proses ini digunakan *GoogleSheet* dan *ad-ons* untuk membersihkan dan merapikan data. Hasil *text preprocessing* dipaparkan di Tabel 2.

Tabel 2. Hasil *Text Preprocessing*

Komentar	Hasil Preprocessing
krispi banget, ngomong sama mahluk ghoib	krispi banget ngomong sama mahluk ghoib
Yang kesini gara gara YouTube rewind like 🤔🤔	yang kesini gara gara youtube rewind like
Menteri taunya cuma duit doang n korupsi Prestasi nihil	menteri taunya cuma duit doang n korupsi prestasi nihil
BPJS doesn't cover that bum, bro.	bpjs doesnt cover that burn bro
Semoga Indonesia lebih baik dengan adanya orang orang seperti mba nana untuk kritik yang membangun □ pertanyaan bagus loh, jdi penasaran bgt jawabannya, sayang gabisa hadir	semoga indonesia lebih baik dengan adanya orang orang seperti mba nana untuk kritik yang membangun pertanyaan bagus loh jadi penasaran banget jawabannya sayang gabisa hadir
Hahahaha... 🤔🤔🤔	hahahaha
Gile yang komen 58 ribu.. luar biasa mba nana	gile yang komen ribu luar biasa mba nana
Plot twist : Terawan is just the chair's name	plot twist terawan is just the chairs name
pak?? kemana pak meninggoy	pak kemana pak meninggoy

### 4.3. Klasifikasi Naïve Bayes

Proses *training* dan proses *testing* adalah tahapan dari klasifikasi *Naïve Bayes*. Proses *training* merupakan proses yang digunakan sebagai model analisis sentimen terhadap data *testing* atau proses *testing*.

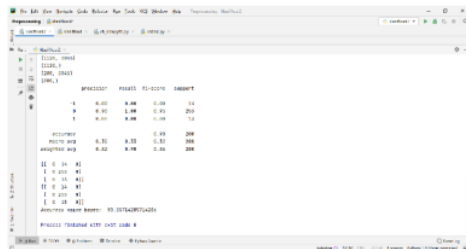
Penelitian ini dilakukan dengan 1400 data komentar yang telah melalui tahap *preprocessing* dan *labelling*. Data tersebut dibagi menjadi data *training* sebanyak 1120 data dan data *testing* sebanyak 280 data. Untuk memodelkan proses klasifikasi *Naïve Bayes* digunakan data *training*, sedangkan data *testing* akan digunakan untuk menguji akurasi dari model yang dibentuk[15]. Proses klasifikasi *Naïve Bayes* dilakukan dengan Bahasa pemrograman python dengan *library scikit learn*.

Hasil klasifikasi menunjukkan data negatif data positif dan netral. Algoritma *Naïve Bayes* dalam penelitian ini menghasilkan akurasi sebesar 90,36%. Berikut ini di Tabel 3 dipaparkan hasil klasifikasi.

Tabel 3. Hasil Klasifikasi

Sentimen	Polaritas
Positif	78
Negatif	90
Netral	1232

Hasil klasifikasi juga dapat dilihat dari jalannya program di Gambar 3 sebagai berikut:



Gambar 3. Hasil Klasifikasi dan Confussion Matrix

## 5. KESIMPULAN

Hasil dan pembahasan mengenai analisis sentimen pengguna Youtube terhadap tayangan #MataNajwaMenantiTerawan dengan metode *Naive Bayes Classifier* diperoleh kesimpulan sebagai berikut:

1. Metode *Naive Bayes* Mampu Melakukan Klasifikasi komentar secara otomatis dengan bahasa pemrograman python.
2. Analisis sentimen terhadap data komentar Youtube terhadap tayangan #MataNajwaMenantiTerawan dapat dilakukan dengan metode *Naive Bayes* dengan nilai akurasi 90.36%.

## 6. DAFTAR PUSTAKA

- [1] Z. A. Achmad, "Perbandingan Sistem Pers," *Lutfansah Mediat.*, 2014, [Online]. Available: [https://www.researchgate.net/profile/Zainal\\_Achmad/publication/326436775\\_Perbandingan\\_Sistem\\_Pers\\_dan\\_Sistem\\_Pers\\_di\\_Indonesia/links/5b4d5a86a6fdcc8dae2468ea/Perbandingan-Sistem-Pers-dan-Sistem-Pers-di-Indonesia.pdf](https://www.researchgate.net/profile/Zainal_Achmad/publication/326436775_Perbandingan_Sistem_Pers_dan_Sistem_Pers_di_Indonesia/links/5b4d5a86a6fdcc8dae2468ea/Perbandingan-Sistem-Pers-dan-Sistem-Pers-di-Indonesia.pdf).
- [2] D. Abul-Fottouh, M. Y. Song, and A. Gruzd, "Examining algorithmic biases in YouTube's recommendations of vaccine videos," *Int. J. Med. Inform.*, vol. 140, no. April, p. 104175, 2020, doi: 10.1016/j.ijmedinf.2020.104175.
- [3] G. A. Buntoro, "Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter," vol. 9 no. 1, pp. 32–41, 2017.
- [4] M. Christianto, J. Andjarwirawan, and A. Tjondrowiguno, "Aplikasi analisa sentimen pada komentar berbahasa Indonesia dalam objek video di website YouTube menggunakan metode *Naive Bayes classifier*," *J. Infra*, vol. 8.1, pp. 255–259, 2020.
- [5] P. Y. Saputra, D. H. Subhi, and F. Z. A. Winatama, "Implementasi Sentimen Analisis Komentar Channel Video Pelayanan Pemerintah Di Youtube Menggunakan Algoritma *Naive Bayes*," *J. Inform. Polinema*, vol. 5, no. 4, pp. 209–213, 2019, doi: 10.33795/jip.v5i4.259.
- [6] I. Algoritma, N. Bayes, and D. A. N. D. Tree, "Perbandingan algoritma *naive bayes*, *svm*, dan *decision tree* untuk klasifikasi sms spam," vol. 05, no. 5, pp. 167–174, 2020.
- [7] Samsir, Ambiyar, U. Verawardina, F. Edi, and R. Watrianthos, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode *Naive Bayes*," *J. Media Inform. Budidarma*, vol. 5, pp. 157–163, 2021, doi: 10.30865/mib.v5i1.2604.
- [8] I. Nurmawati, E. Supriyati, and T. Listyorini, "Analisis Sentimen Terhadap Penggemar Kpop Di Media Sosial Twitter Menggunakan *Naive Bayes* (Studi Kasus Penggemar Grup Bts)," *JIEET (Journal Inf. Eng. Educ. Technol.*, vol. 4, no. 2, pp. 86–89, 2020, [Online]. Available: <https://journal.unesa.ac.id/index.php/jieet/article/view/11522>.
- [9] Balya, "Analisis Sentimen Pengguna Youtube Di Indonesia Pada Review Smartphone Menggunakan *Naive Bayes*," pp. 4–16, 2003.
- [10] J. F. Sánchez-Rada and C. A. Iglesias, "Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison," *Inf. Fusion*, vol. 52, no. May, pp. 344–356, 2019, doi: 10.1016/j.inffus.2019.05.003.
- [11] Y. Cahyono and S. Saprudin, "Analisis Sentiment Tweets Berbahasa Sunda Menggunakan *Naive Bayes Classifier* dengan Seleksi Feature Chi Squared Statistic," *J. Inform. Univ. Pamulang*, vol. 4, no. 3, p. 87, 2019, doi: 10.32493/informatika.v4i3.3186.
- [12] B. Haryanto, Y. Ruldeviyani, F. Rohman, T. N. Julius Dimas, R. Magdalena, and F. Muhamad Yasil, "Facebook analysis of community sentiment on 2019 Indonesian presidential candidates from Facebook opinion data," *Procedia Comput. Sci.*, vol. 161, pp. 715–722, 2019, doi: 10.1016/j.procs.2019.11.175.
- [13] N. Yusuf, E. Supriyati, and T. Listyorini, "Analisis Sentimen Mengenai Layanan Provider Indihome Berdasarkan Pendapat Pelanggan Melalui Media Sosial Twitter dengan Metode *Naive Bayes Classifier*," vol. 6, pp. 75–78, 2020.
- [14] L. L. Dhande and G. K. Patnaik, "Analyzing Sentiment of Movie Review Data using *Naive Bayes Neural Classifier*," *Int. J. Emerg. Trends Technol. Comput. Sci.*, vol. 3, no. 4, pp. 313–340, 2014.
- [15] A. Indrawati, "Penerapan Teknik Kombinasi *Oversampling* Dan *Undersampling* Untuk Mengatasi Permasalahan *Imbalanced*

Dataset,” *22 O (Jurnal Inform. dan Komputer)*, vol. 4, no. 1, pp. 38–43, 2021, doi: 10.33387/jiko.v4i1.2561.

## ORIGINALITY REPORT

**20%**  
SIMILARITY INDEX

**18%**  
INTERNET SOURCES

**10%**  
PUBLICATIONS

**8%**  
STUDENT PAPERS

## PRIMARY SOURCES

<b>1</b>	<a href="https://repositori.usu.ac.id">repositori.usu.ac.id</a> Internet Source	<b>2%</b>
<b>2</b>	Submitted to Universitas Bunda Mulia Student Paper	<b>2%</b>
<b>3</b>	Ahmed Omar, Tarek M. Mahmoud, Tarek Abd-El-Hafeez, Ahmed Mahfouz. "Multi-label Arabic text classification in Online Social Networks", Information Systems, 2021 Publication	<b>1%</b>
<b>4</b>	<a href="http://ejournal.unkhair.ac.id">ejournal.unkhair.ac.id</a> Internet Source	<b>1%</b>
<b>5</b>	<a href="http://ejurnal.univalabuhanbatu.ac.id">ejurnal.univalabuhanbatu.ac.id</a> Internet Source	<b>1%</b>
<b>6</b>	Sari Widya Sihwi, Insan Prasetya Jati, Rini Anggrainingsih. "Twitter Sentiment Analysis of Movie Reviews Using Information Gain and Naïve Bayes Classifier", 2018 International Seminar on Application for Technology of Information and Communication, 2018 Publication	<b>1%</b>

7	<a href="http://journal.unesa.ac.id">journal.unesa.ac.id</a> Internet Source	1 %
8	<a href="http://ijcit.com">ijcit.com</a> Internet Source	1 %
9	<a href="http://www.warse.org">www.warse.org</a> Internet Source	1 %
10	<a href="http://ejournals.umn.ac.id">ejournals.umn.ac.id</a> Internet Source	1 %
11	<a href="http://ejurnal.stmik-budidarma.ac.id">ejurnal.stmik-budidarma.ac.id</a> Internet Source	1 %
12	<a href="http://www.teknikaftui.net">www.teknikaftui.net</a> Internet Source	1 %
13	<a href="http://www.researchgate.net">www.researchgate.net</a> Internet Source	1 %
14	"Abstracts", Basic & Clinical Pharmacology & Toxicology, 2021 Publication	1 %
15	Submitted to Konsorsium Perguruan Tinggi Swasta Indonesia II Student Paper	1 %
16	<a href="http://journal.uii.ac.id">journal.uii.ac.id</a> Internet Source	1 %
17	<a href="http://ecampus.pelitabangsa.ac.id">ecampus.pelitabangsa.ac.id</a> Internet Source	<1 %

18

Ni Made Ayu Juli Astari, Dewa Gede Hendra Divayana, Gede Indrawan. "Analisis Sentimen Dokumen Twitter Mengenai Dampak Virus Corona Menggunakan Metode Naive Bayes Classifier", Jurnal Sistem dan Informatika (JSI), 2020

Publication

&lt;1 %

19

ppjp.ulm.ac.id

Internet Source

&lt;1 %

20

text-id.123dok.com

Internet Source

&lt;1 %

21

Khan, Farhan Hassan, Usman Qamar, and Saba Bashir. "SentiMI: Introducing point-wise mutual information with SentiWordNet to improve sentiment polarity detection", Applied Soft Computing, 2016.

Publication

&lt;1 %

22

Muhammad Ridho Matondang, Muhammad Ridwan Lubis, Heru Satria Tambunan. "Analisis Data mining dengan Metode C.45 pada Klasifikasi Kenaikan Rata-Rata Volume Perikanan Tangkap", Brahmana : Jurnal Penerapan Kecerdasan Buatan, 2021

Publication

&lt;1 %

23

ejournal.unitomo.ac.id

Internet Source

&lt;1 %

[jurnal.ugm.ac.id](http://jurnal.ugm.ac.id)

24

Internet Source

<1 %

25

repository.ummat.ac.id  
Internet Source

<1 %

26

Deena Abul-Fottouh, Melodie Yunju Song, Anatoliy Gruzd. "Examining algorithmic biases in YouTube's recommendations of vaccine videos", International Journal of Medical Informatics, 2020  
Publication

<1 %

Exclude quotes Off  
Exclude bibliography Off

Exclude matches Off