

**Jutisi:** Jurnal Ilmiah Teknik Informatika dan Sistem Informasi  
 Jl. Ahmad Yani, K.M. 33,5 - Kampus STMIK Banjarbaru  
 Loktabat – Banjarbaru (Tlp. 0511 4782881), e-mail: puslit.stmikbjb@gmail.com  
 e-ISSN: [2685-0893](#)  
 p-ISSN: 2089-3787

## Analisis Sentimen Warganet Terhadap KTT G20 Bali Menggunakan Algoritme Naïve Bayes

**Abdul Halim Anshor**

Program Studi Teknik Informatika, Universitas Pelita Bangsa  
 Jl. Inspeksi Kalimalang No.9, Cibatu, Bekasi, Indonesia  
 Email Corresponding Author: [abdulhalimanshor@pelitabangsa.ac.id](mailto:abdulhalimanshor@pelitabangsa.ac.id)

### Abstract

*As an effort to evaluate the implementation of the G20 Bali High Level Conference (Summit) activities, it is appropriate for the Indonesian government to receive input from various parties, including analyzing public responses (sentiments) regarding the implementation of these activities, so that they can obtain input for improving similar activities in the future. This paper presents an analysis of the sentiments of Twitter users towards the implementation of the Bali G20 Summit. The analysis aims to classify community responses into positive and negative categories using the Naïve Bayes Classifier (NBC) algorithm. The analysis process was carried out by crawling Twitter data, preprocessing, classification using the Naïve Bayes algorithm, validating results, and evaluating. The results of the Naïve Bayes performance test show an accuracy value of 94.55%, 100% precision and 89.10% recall. These results are expected to be information supporting decisions for the government in making policies related to G20 cooperation in the future.*

**Keywords:** *High Level Conference; G20, Classification; Netizens' Responses; Twitter*

### Abstrak

Sebagai upaya mengevaluasi pelaksanaan kegiatan Konferensi Tingkat Tinggi (KTT) G20 Bali, sudah selayaknya pemerintah Indonesia menerima masukan-masuk dari berbagai pihak, termasuk menganalisis tanggapan (sentimen) masyarakat mengenai pelaksanaan kegiatan tersebut, sehingga dapat memperoleh masukan untuk penyempurnaan kegiatan serupa di masa mendatang. Paper ini menyajikan analisa sentimen masyarakat pengguna *Twitter* terhadap pelaksanaan KTT G20 Bali. Analisa bertujuan untuk mengklasifikasikan tanggapan masyarakat ke dalam kategori positif dan negatif dengan menggunakan algoritme *Naïve Bayes Classifier* (NBC). Proses analisa dilakukan dengan tahapan *crawling* data *twitter*, *Preprocessing*, Klasifikasi menggunakan Algoritme *Naïve Bayes*, validasi hasil, dan evaluasi. Hasil pengujian kinerja *Naïve Bayes* menunjukkan nilai *Accuracy* sebesar 94,55%, *precision* 100%, dan *recall* 89,10%. Hasil ini diharapkan dapat menjadi informasi pendukung keputusan bagi pemerintah dalam mengambil kebijakan terkait kerjasama G20 dimasa yang akan datang.

**Kata kunci:** *Konferensi Tingkat Tinggi; G20; Klasifikasi; Tanggapan Warganet; Twitter*

### 1. Pendahuluan

G20 adalah bentuk kerjasama antar negara G20 berupa forum kerja sama multilateral yang terdiri dari 19 negara utama dan Uni Eropa (EU). G20 merepresentasikan lebih dari 60% populasi bumi, 75% perdagangan global, dan 80% PDB dunia. Anggota G20 terdiri dari Afrika Selatan, Amerika Serikat, Arab Saudi, Argentina, Australia, Brasil, India, Indonesia, Inggris, Italia, Jepang, Jerman, Kanada, Meksiko, Republik Korea, Rusia, Perancis, Cina, Turki, dan Uni Eropa. Indonesia sebagai pemegang presidensi G20 telah menyelenggarakan konferensi tingkat tinggi (KTT) G20 yang ke 17 di Bali pada tanggal 10 November sampai dengan tanggal 17 November 2022, dengan tema *Recover together, Recover Stronger*. Dengan tema tersebut, Indonesia mengajak seluruh dunia untuk bahu-membahu, saling mendukung untuk pulih bersama serta tumbuh lebih kuat dan berkelanjutan. Presidensi G20 Indonesia tentu saja memberikan dampak pada pemulihan perekonomian Indonesia Pelaksanaan KTT G20 di Bali mendapat berbagai respon atau tanggapan dari masyarakat luas, baik berupa respon positif

maupun respon negatif. Masyarakat pada umumnya menyampaikan respon mereka melalui media sosial, salah satu di antaranya adalah media *Twitter*, yang merupakan media sosial yang banyak digunakan dan luas jangkauan penyebarannya.

Sebagai upaya mengevaluasi pelaksanaan kegiatan KTT G20 di Bali tersebut, sudah selayaknya pemerintah menerima masukan-masukan dari berbagai pihak, termasuk menganalisis tanggapan (sentimen) masyarakat mengenai pelaksanaan kegiatan, sehingga dapat menjadi masukan untuk penyempurnaan kegiatan serupa di masa-masa mendatang. Agar pemrosesan dan penyajian hasil analisis tanggapan masyarakat dapat berjalan efektif dan efisien, diperlukan suatu model komputasi tertentu untuk memperoleh hasil yang akurat serta mempercepat proses penyajian hasil, dibandingkan jika itu dilakukan secara manual [1, 2].

Dalam teknologi *Data Mining*, terdapat berbagai model algoritme komputasi yang dapat digunakan untuk menganalisis data, salah satu diantaranya adalah algoritme *Naïve Bayes* [3]. Algoritme *Naïve Bayes* merupakan algoritme yang menerapkan teknik *supervised* klasifikasi objek di masa depan dengan menetapkan label kelas ke *instance* menggunakan probabilitas bersyarat [4]. Algoritme *Naïve Bayes* telah diuji dalam menganalisis data pada berbagai bidang, seperti dalam bidang kesehatan [5, 6], bidang-bidang yang berkaitan dengan kebijakan publik [7, 8], serta bidang-bidang bisnis lainnya [9 – 11].

Paper ini menguji penggunaan algoritma *Naïve Bayes* dalam menganalisis sentimen warganet pengguna *Twitter* terhadap pelaksanaan KTT G20 di Bali. Hasil kajian diharapkan dapat dijadikan sebagai bahan informasi pendukung keputusan bagi pemerintah dalam mengambil kebijakan terkait kerjasama G20 untuk kemajuan bangsa Indonesia dimasa yang akan datang.

## 2. Tinjauan Pustaka

Beberapa penelitian terdahulu yang berkaitan dengan analisis sentimen masyarakat pengguna media sosial telah dilakukan. Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial *Twitter* dilakukan oleh Nurjanah, Perdana, dan Fauzi [12]. Metode yang digunakan dalam analisis tersebut adalah *K-Nearest Neighbor (K-NN)*, dengan menambahkan fitur pembobotan jumlah retweet (non-tekstual). Pembobotan tekstual hasil dari klasifikasi algoritma K-NN dan pembobotan non-tekstual dari pembobotan jumlah retweet akan digabungkan menggunakan nilai konstanta tertentu ( $\alpha$  dan  $\beta$ ) untuk menghasilkan sentimen positif dan negatif. Data yang digunakan berupa opini masyarakat terhadap tayangan televisi pada *twitter* sejumlah 400. Hasil pengujian akurasi metode menggunakan pembobotan tekstual diperoleh 82,50%, menggunakan pembobotan non-tekstual 60%, dan menggunakan penggabungan keduanya 83,33%.

Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial *Twitter* dilakukan oleh Rachman dan Pramana [13]. Analisis memanfaatkan data yang bersumber dari media sosial *twitter*. Penelitian tersebut menganalisis respon masyarakat terhadap wacana vaksinasi dengan cara mengklasifikasikan respon masyarakat ke dalam respon positif dan negatif. Selanjutnya dilakukan pengelompokan opini masyarakat menggunakan metode *Latent Dirichlet Allocation (LDA)* untuk mengetahui topik pembicaraan yang sering dibahas oleh masyarakat terkait dengan wacana vaksinasi tersebut. Hasil analisis pada uji coba tersebut menunjukkan masyarakat lebih banyak memberikan respon positif, yaitu 30%, dibandingkan dengan respon negatifnya yang hanya 26%.

Penelitian yang dilakukan oleh Sya'bani, Fais, Enri, dan Padilah dengan judul "Analisis Sentimen Terhadap Bakal Calon Presiden 2024 Dengan Algoritme *Naïve Bayes*." Hasil penelitian menyimpulkan bahwa warganet memiliki sentimen positif terhadap setiap tokoh bakal calon presiden yang akan datang. Kemudian untuk hasil evaluasi dari algoritme *naïve bayes* yang didapatkan dari dataset pertama adalah 73,68 akurasi dan AUC 0,74 pada fold ke-7, dataset kedua adalah 71,43 untuk akurasi dan AUC 1,0 pada fold ke – 5, untuk dataset ketiga nilai akurasi yang didapat 60% dan AUC 0,92 pada fold ke-1, dan untuk dataset terakhir nilai akurasi yang didapatkan adalah 62,5% dan AUC 0,65 pada fold ke-3 [14].

Pada penelitian yang dilakukan oleh Mimboro dengan judul "Analisis Sentimen Twitter terhadap PILKADA kota Medan menggunakan metode *Naïve Bayes*". Analisis Sentimen dilakukan dengan klasifikasi *tweet* yang mengandung sentimen Netizen terhadap Penyelenggaraan Pilkada Kota Medan 2020. Metode klasifikasi yang digunakan dalam penelitian ini adalah metode *Naïve Bayes* yang dikombinasikan dengan ekstraksi fitur TF-IDF. Uji validitas yang diterapkan pada penelitian ini menggunakan matriks konfusi. Dengan

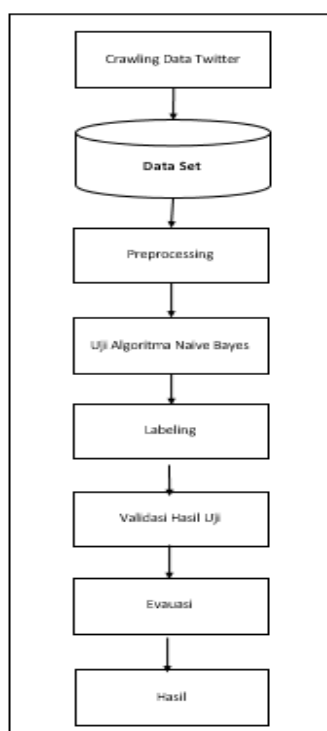
fiturtf-idf ekstraksi dan metode Naïve Bayes akan dapat secara otomatis mengklasifikasikan analisis sentiment dengan hasil akurasi 76,00% [15].

Penelitian yang dilakukan oleh Nurhazizah, Ichsan, dan Widiyanesti yang berjudul "Analisis Sentimen Dan Jaringan Sosial Pada Penyebaran Informasi Vaksinasi Di Twitter." melakukan analisis sentimen terhadap opini terkait vaksin yang beredar di twitter serta melakukan analisis jaringan sosial (SNA) yang terbentuk untuk mengetahui aktor yang paling berperan dalam penyebaran informasi mengenai vaksin COVID-19. Penelitian ini menggunakan metode klasifikasi dengan algoritma naïve bayes dan metode Social Network Analysis (SNA). Hasil dari penelitian ini menunjukkan bahwa 92% sentimen pengguna twitter adalah positif terhadap vaksin COVID-19 dan aktor yang paling berperan dalam penyebaran informasi adalah akun @jokowi [16].

Penelitian yang kami lakukan berbeda dengan penelitian [12] dan [13], yaitu menggunakan algoritme yang berbeda dalam proses analisis sentimen opini masyarakat pengguna soaial media. Namun demikian, penelitian kami memiliki kesamaan dengan penelitian [14] – [16], yaitu menggunakan algoritme *Naïve Bayes* dalam menganalisis sentimen opini masyarakat pengguna soaial. Perbedaan hany a terdapat pada Objek Kajian. Pada penelitian kami, objek yang dikaji berkaitan dengan sentimen masyarakat pengguna media sosial *Twitter* terhadap pelaksanaan KTT G20 di Bali. Dalam paper kami, pengklasifikasian data dilakukan pada data sentimen positif dan negatif berupa *tweets* berbahasa Indonesia.

### 3. Metodologi

Objek Penelitian yang merupakan permasalahan yang dianalisis dan dibahas dalam paper ini adalah masalah pelaksanaan KTT G20 di Bali. Metode yang digunakan dalam pengklasifikasian tanggapan masyarakat positif dan negatif adalah menggunakan algoritme *Naïve bayes*. Pengolahan dan analisis data menggunakan *tools Rapid minner*. Data set yang digunakan berupa data *tweets* yang berisi tema KTT G20 di bali sebanyak 1500 sampel. Pengujian akurasi kinerja metode melalui uji presisi dan *recall* dari sentimen warganet *Twitter*.



Gambar 1. Prosedur Analisis Data Berbasis Algoritme Naïve Bayes

Adapun data set yang digunakan dalam proses analisis dikelompokkan menjadi data training dan data testing. Pada langkah awal akan dilakukan preprocessing data selanjutnya. memberikan sentimen (labeling) terhadap *review text dataset*, setelah itu data yang telah diberi

sentimen dikumpulkan antara sentimen positif dan sentimen negatif menjadi data training dan data testing, selanjutnya akan dilakukan model klasifikasi *Naive Bayes* dan mendapatkan hasil *Confusion Matrix*, selanjutnya dihitung hasil akurasi kinerja algoritma *Naive Bayes* tersebut.

Data yang sudah diberi label sentiment positif dan sentiment negative selanjutnya akan diuji dengan algoritme *Naive Bayes*. *Naive Bayes* sebagai algoritme klasifikasi menghitung probabilitas untuk mengurangi data yang kompleks menjadi beberapa probabilitas sederhana dan metode ini membutuhkan beberapa data yang dilatih untuk menentukan parameter yang dibutuhkan dalam proses klasifikasian. Bentuk formula dari *Naive Bayes* sebagai berikut:

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

Keterangan:

X = Data dengan kelas yang belum diketahui

H = Hipotesa data X merupakan suatu kelas spesifik

P(H|X) = Probabilitas hipotesis H berdasarkan kondisi X (*posterior probability*)

P(H) = Probabilitas hipotesis H (*prior probability*)

Data dibagi dalam komposisi 60% data training dan 40% data testing. Akurasi yang dihitung adalah akurasi data testing dengan menggunakan formulasi *confusion matrix* seperti pada Tabel 1 dan Formula 2.

Tabel 1 *Confusion Matrix*

		<i>True Class</i>	
		Positif	Negatif
<i>Predicted Class</i>	Positif	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>
	Negatif	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>

Rumus perhitungan *Confusion Matrix*:

$$Accuracy (A) = (TP + TN) / ((TP + TN + FP + FN)) \dots\dots (2)$$

#### 4. Hasil dan Pembahasan

Pada penelitian ini jumlah dataset yang digunakan sebanyak 396 data tweet yang terdiri dari tweet positif dan tweet negatif. Proses crawling data dengan kata kunci KTT G20 di Bali sebagai objek utama penelitiannya. Proses pengumpulan data dilakukan dengan proses crawling data yang menggunakan perangkat lunak RapidMiner. Jumlah data yang diperoleh dari proses tersebut sebanyak 1500 data. Setelah itu dilakukan proses preprocessing atau seleksi data, dataset yang diperoleh sebanyak 396 data. Untuk lebih jelasnya dapat dilihat pada Tabel 1.

Tabel 1. Dataset

Kata Kunci Twitter	Data Awal	Data Akhir	Positif	Negatif
KTT G20 di Bali	1500	396	235	161

#### 4.1 Pemrosesan data Berbasis Algoritma *Naive Bayes*.

##### 1) *Crawling Data*

Pada proses *Crawling* data menggunakan *RapidMiner* dengan menggunakan operator *connection twitter*. *Search twitter* kemudian disimpan dalam file.csv. Parameter yang digunakan sebagai berikut, key words yang dipakai adalah "KTT G20 di Bali", jumlah data yang diambil sebanyak 1500 record data. Data yang diambil adalah tweets atau komentar yang berbahasa indonesia saja. Berikut dibawah ini adalah proses crawling twitter beserta data hasilnya:



Gambar 2 Process *Crawling* Data

2) *Seleksi Data*

Setelah proses crawling proses selanjutnya adalah memfilter kolom yang akan di pakai saja yaitu kolom text. Pada kolom ini berisi komentar yang mengandung sentiment yang akan dihitung akurasinya, seperti gambar 3.



Gambar 3. Parameter *Select Attribute*

3) *Cleaning*

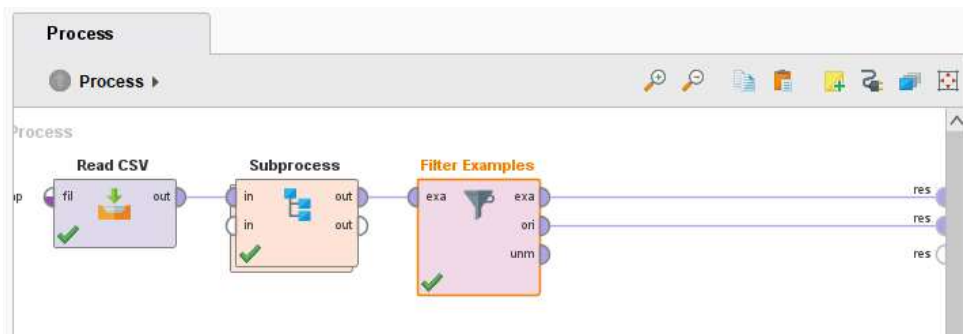
Data hasil crawling twitter masih terdapat banyak item yang tidak diperlukan dalam proses pengolahan data analisa data untuk itu item-item tersebut perlu dihapus atau dihilangkan. Setelah proses *Cleaning*, diperoleh informasi seperti beriku:



Gambar 4. Informasi Yang Diperoleh Setelah Proses *Cleaning*

4) *Filter Examples*

Operator *Filter Examples* ini akan digunakan untuk mengambil data yang ada isinya saja dan yang tidak memiliki isi atau kosong akan dibuang. Proses *Filter* disajikan pada Gambar 5 dan hasil proses seperti disajikan pada Gambar 6.



Gambar 5. Proses *Filter Examples Not Missing*

Row No.	Text
1	Hari ini saya minta dibentuk gugus tugas menindaklanjuti berbagai kesepakatan tersebut sifto/CN2s88FcV
2	Hanrisoon Gelaran KTT G20 di Bali btp waitu lalu memberikan makna dan kesan tersendiri bagi Indonesia sebagai Tuan Rumah Menteri BUM...
3	jokowi KTT G20 di Bali menghasilkan kesepakatan setidaknya 225 proyek multilateral senilai USD238 miliar dan 140 proyek bilateral seni...
4	Menteri Keuangan Sri Mulyani Indrawati membagikan momennya saat duduk bersampingan bersama Menteri Pertahanan Menteri Prabowo Subianto saat h...
5	Menteri Keuangan Sri Mulyani Indrawati membagikan momennya saat duduk bersampingan bersama Menteri Pertahanan Menteri Prabowo Subianto saat h...
6	PT Wasita Karya turut mendukung persiapan penyelenggaraan Presidensi KTT G20 di Bali pada 15-16 November 2022
7	WasitaDukungG20
8	G20Indonesia
9	WasitaKarya
10	BUMHUntukIndonesia sifto/Cd0t6RFFh
11	sifto/dBYEEep3X
12	PT Wasita Karya turut mendukung persiapan penyelenggaraan Presidensi KTT G20 di Bali pada 15-16 November 2022.Untuk itu Wasita telah menyelesaikan...
13	Bonni.Jalani Tidak hanya di Bali Wasita Karya juga turut mendukung penyelenggaraan Presidensi KTT G20 dengan diselesainya proyek...
14	BangkitBersamaET sifto/M5Sm45jeGp
15	Kemendagri 1 Presiden Jokowi saat memimpin rapat terbatas tentang evaluasi pelaksanaan KTT G20 di Kantor Presiden Jakarta pada...

Gambar 6. Informasi Sesudah Difilter Dengan *Filter Examples*

### 5) *Remove Duplikat*

Setelah proses Filter, selanjutnya dilakukan *Remove Duplicate* untuk menghilangkan *tweet* yang sama. Hasilnya seperti disajikan pada Gambar 7.

Row No.	Text
1	Hari ini saya minta dibentuk gugus tugas menindaklanjuti berbagai kesepakatan tersebut sifto/CN2s88FcV
2	Hanrisoon Gelaran KTT G20 di Bali btp waitu lalu memberikan makna dan kesan tersendiri bagi Indonesia sebagai Tuan Rumah Menteri BUM...
3	jokowi KTT G20 di Bali menghasilkan kesepakatan setidaknya 225 proyek multilateral senilai USD238 miliar dan 140 proyek bilateral seni...
4	Menteri Keuangan Sri Mulyani Indrawati membagikan momennya saat duduk bersampingan bersama Menteri Pertahanan Menteri Prabowo Subianto saat h...
5	Menteri Keuangan Sri Mulyani Indrawati membagikan momennya saat duduk bersampingan bersama Menteri Pertahanan Menteri Prabowo Subianto saat h...
6	PT Wasita Karya turut mendukung persiapan penyelenggaraan Presidensi KTT G20 di Bali pada 15-16 November 2022

Gambar 7. Data *Duplicate*

Pada Gambar 7 terlihat data nomor 4 dan data nomor 5 adalah data yang duplikat, sehingga perlu dilakukan proses *Remove Duplicate*.

### 6) *Pelabelan Manual*

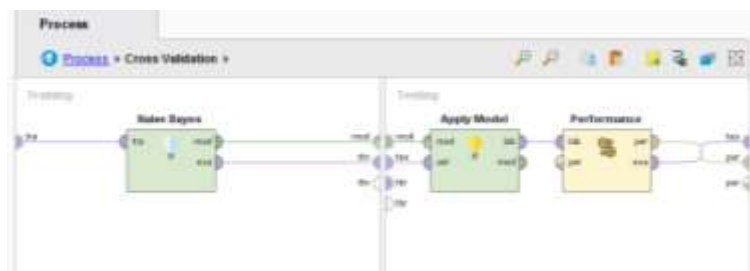
Setelah proses *preprocessing* kemudian dilakukan pelabelan sentimen pada setiap kalimat secara *manual*. Pelabelan yang diberikan berkaitan dengan apakah *tweets* termasuk sentimen positif atau sentimen negatif. Data akan dipecah menjadi 2 bagian yaitu data *training* dan data *testing*. Data yang akan diberi sentimen secara *manual* adalah data *training* sebanyak 235 data dan sisanya adalah 161 data untuk data *testing* yang akan diuji pakai algoritma *Naive Bayes*. Hasil proses seperti disajikan pada Gambar 8.

Row No.	Sentimen	Text
1	Positif	... mofurn jepang asli Indonesia aku jugaaaa hanaui lgcream #tt g20 di bali john wick na...
2	Positif	[cm] Hasil KTT G20 di Bali tentang lingkungan itu membahas apa aja ya pas baca rangkumannya tipe cuman "adaptasi dengan p...
3	Positif	1 Presiden Jokowi saat memimpin rapat terbatas tentang evaluasi pelaksanaan KTT G20 di Kantor Presiden Jakarta pada Senin 28...
4	Positif	1 Presideni G20 Indonesia dengan puncak acara yakni KTT G20 akan diselenggarakan di Bali pada 15-16 November 2022 Acara L...
5	Positif	3 Pertemuan yang berlangsung selama kurang lebih 30 menit ini merupakan pertemuan pertama kedua pemimpin setelah perhelat...
6	Positif	6 Masduki menegaskan bahwa pertemuan kedua pemimpin ini menandakan hubungan harmonis diantara keduanya Sebab seba...
7	Positif	662 untuk lebih digunakan untuk menunjang kegiatan para delegasi negara selama KTT G20 di Bali
8	Positif	8Yunita_ ini merupakan sebuah respon positif yg sangat berarti karena IndonesiaIndonesia telah sukses menuntaskan perhelatan...
9	Positif	AbangDatus Direktur Utama Pertamina Putra Niaga Alfan Haidilun menyatakan tawaran Bahan Bakar Minyak BBM dan Liquefied P...
10	Positif	AbangDatus Pertamina Putra Niaga mendukung Pemerintah dalam penggunaan energi bersih pada KTT G20 di Bali 15-16 Novemb...
11	Positif	Adopti Penuh Bali G20 Leaders' D...
12	Positif	Adopti Penuh Bali G20 Leaders' Declaration
13	Positif	Alpa Enkand dan Bripta Fady mendapat Penghargaan dari Kakor Lantas Pom yang diberikan oleh Kapotes Luwu Timur aitol...
14	Positif	Aiyahda1 Pak Jokowi kare telah membuat harum nama Indonesia dunia km telah berhasil membawa pesan damai dan suksa...
15	Positif	atymny didalam KTT G20 yg di selenggarakan di bali beberapa waktu lalu Pertamina berkomitmen untuk penggunaan energi rama...

Gambar 8. Proses Pelabelan Manual terhadap Suatu Pernyataan Sentimen

#### 4.2 Pengujian Kinerja Algoritma *Naive Bayes*.

Peneliti membagi membagi data menjadi dua yaitu data *training* dan data *testing*. Adapun Item yang akan dihitung tingkat *accuracy*, *precision*, dan *recall* terhadap data *testing* menggunakan Rapidminer untuk mengetahui seberapa akurat hasil pengujian data tersebut. Untuk mendapatkan hasil pengujian yang optimal peneliti membagi komposisi 60% data training dan 40% data testing.



Gambar 4. Proses Pengujian Data *testing* 40% dengan Algoritma *Naive Bayes*

accuracy: 94.5% ← 50% (micro-average: 94.5%)

	pred Positif	pred Negatif	class precision
pred Positif	138	0	100.0%
pred Negatif	0	166	99.4%
class recall	89.1%	100.0%	

Gambar 9. Hasil Pengujian Data *Testing* 40% dengan Algoritma *Naive Bayes*

##### 1. *Accuracy*

Akurasi adalah tingkat kedekatan antara hasil prediksi dengan hasil fakta, prediksi adalah tingkat ketetapan antara informasi yang diminta oleh *user* dengan jawaban yang diberikan oleh sistem. Setelah data sudah terbagi kemudian ditentukan nilai prediksinya. Prediksi dilakukan dengan menggunakan data *testing* 40%.

Adapun hasil *accuracy* diperoleh dengan cara menjumlah data *true positive* (TP) ditambah dengan *true negative* (TN) dibagi dengan total jumlah data *testing* 40%. Berikut cara perhitungan manual dari *accuracy*:

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \\
 &= \frac{139+156}{139+156+0+17} \times 100\% \\
 &= \frac{295}{312} \times 100\% \\
 &= 94.55\%
 \end{aligned}$$

Dari hasil pengujian data sentimen negatif yang terdeteksi negatif adalah 156 dan yang salah adalah 0, sedangkan data sentimen positif yang terdeteksi positif adalah 139 dan yang salah adalah 17. Hasil dari *accuracy Naive Bayes* dengan *Rapidminer* adalah 94.55%.

## 2. Precision

Presisi merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Pada pengujian ini *precision* terhadap data *testing* dihitung menggunakan *Rapidminer*. Prediksi dilakukan dengan menggunakan data *testing* 40%.

Adapun hasil *precision* diperoleh dengan cara data *true positive* (TP) dibagi dengan total jumlah data *true positive* ditambah dengan *false positive* (FP). Berikut cara perhitungan manual dari *precision*:

$$\begin{aligned}
 \text{Precision} &= \frac{TP}{TP+FP} \times 100\% \\
 &= \frac{139}{139+0} \times 100\% \\
 &= \frac{139}{139} \times 100\% \\
 &= 100\%
 \end{aligned}$$

Prediksi dari data sentimen positif yang terdeteksi positif adalah 139 dan yang salah adalah 0. Hasil dari *precision Naive Bayes* dengan *Rapidminer* adalah 100.00%.

## 3. Recall

*Recall* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif, adalah pengukuran tingkat keberhasilan pengambilan data yang relevan. Pada pengujian ini akan dihitung *recall* terhadap data *testing* menggunakan *Rapidminer*. Berikut ini hasil *recall* pengujian Algoritma *Naive Bayes* dengan data *testing* 40%:

Prediksi dari data sentimen positif terdeteksi positif adalah 139 sedangkan data sentimen negatif yang salah terdeteksi *negative* adalah 17 adalah 89,10%.

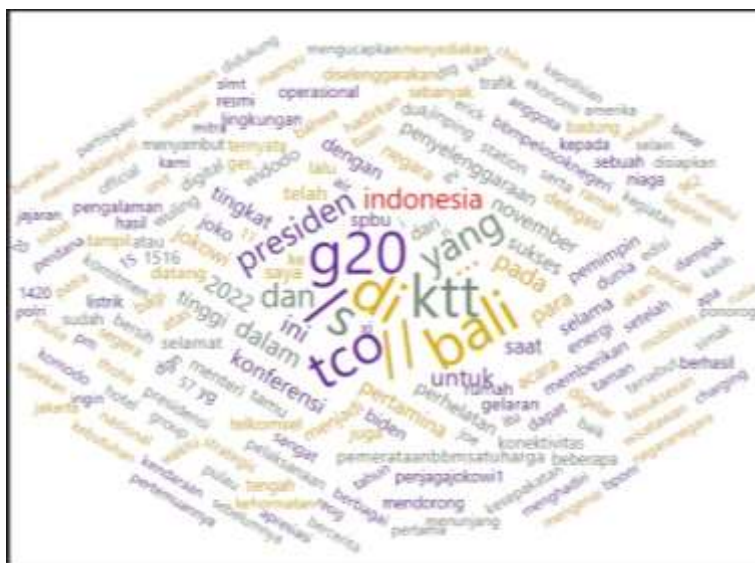
Adapun hasil *recall* diperoleh dengan cara data *true positive* (TP) dibagi dengan total jumlah data *true positive* (TP) ditambah dengan *false negative* (FN). Nilai *recall* diperoleh seperti berikut:

$$\begin{aligned}
 \text{Recall} &= \frac{TP}{TP+FN} \times 100\% \\
 &= \frac{139}{139+17} \times 100\% \\
 &= \frac{139}{156} \times 100\% \\
 &= 89,10\%
 \end{aligned}$$

## 4.3 Pembahasan

Setelah mendapatkan hasil akurasi pada pengujian *Naive Bayes*, kemudian dilakukan proses pembentukan *wordcloud*. *Wordcloud* dilakukan untuk mengetahui frekuensi setiap kata. Untuk menentukan *wordcloud* maka *dataset* akan diubah dengan menggunakan *Rapidminer*. Dalam hasil *wordcloud*, untuk tulisan yang paling besar menandakan bahwa kata tersebut yang paling sering muncul pada teks, begitupun pada ukuran yang lebih kecil menandakan kata tersebut lebih sedikit muncul dalam teks. Pada hasil *wordcloud* dalam penelitian ini, kata yang paling sering muncul adalah ktt, g20, di bali, kemudian dan seterusnya hingga kata yang paling sedikit muncul. Disini diambil minimal kata yang dimunculkan pada *wordcloud* adalah 100 pengulangan kata. Berikut adalah hasil *wordcloud* pada penelitian ini:





Gambar 10. Tampilan Wordcloud

Hasil pengujian akurasi Kinerja algoritma *Naïve Bayes* dalam uji coba penelitian ini mencapai 95%, dengan nilai *precision* sebesar 100% dan *Recall* mencapai 89%. Hasil ini juga menunjukkan bahwa algoritma *Naïve Bayes* merupakan algoritma yang cukup baik digunakan dalam menganalisis sentimen pengguna media sosial terhadap suatu isu-isu tertentu. Temuan ini menguatkan temuan pada penelitian [17] yang memperoleh nilai 92% dan penelitian [18] dengan tingkat akurasi mencapai 93%.

## 5. Simpulan

Dari penelitian yang telah dilakukan maka dapat disimpulkan bahwa pengujian dengan algoritma Naive Bayes menggunakan operator performances menghasilkan tingkat dan hasil yang cukup baik. Nilai accuracy, precision, dan recall tertinggi yang mana mendapatkan hasil dengan nilai accuracy sebanyak 94.55%, precision sebanyak 100%, dan recall sebanyak 89.10% dan menghasilkan hasil opini berbentuk positif terhadap objek penelitian yaitu KTT G20 di bali sebanyak 35% tweets. Hasil ini diharapkan dapat menjadi masukan dalam membangun sistem pendukung keputusan bagi pemerintah dalam mengambil kebijakan terkait suatu pelaksanaan kegiatan kenegaraan di masa mendatang.

## Daftar Referensi

- [1] D. Gunawan, "Evaluasi Performa Pemecahan Database dengan Metode Klasifikasi Pada Data Preprocessing Data mining". *Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika*, vol. 2, no. 1, pp. 10-13, 2016.
- [2] R.A. Cahdriyana, R. Richardo, "Berpikir komputasi dalam pembelajaran matematika". *LITERASI (Jurnal Ilmu Pendidikan)*, vol. 11, no. 1, pp. 50-56, 2020.
- [3] N. L. W. S. R. Ginantra, F.N. Arifah, A.H. Wijaya, R.S. Septarini, N. Ahmad, D. Ardiana, & E.S. Negara, "*Data mining dan penerapan algoritma*". Yayasan Kita Menulis, 2021.
- [4] J. Suntoro, *Data Mining Algoritma dan Implementasi dengan Pemrograman PHP*, Edisi I, Jakarta: Elex Media Komputindo, 2019
- [5] A. Harun, D.P. Ananda, "Analysis of public opinion sentiment about COVID-19 vaccination in Indonesia using Naïve Bayes and Decission Tree Analisa Sentimen opini publik tentang vaksinasi COVID-19 di Indonesia menggunakan Naïve Bayes dan Decission Tree". *Indones. J. Mach. Learn. Comput. Sci*, vol. 1, pp. 58-63, 2021.
- [6] N.R. Wardani, A. Erfina, "Analisis Sentimen Masyarakat Terhadap Layanan Konsultasi Dokter Menggunakan Algoritma Naive Bayes. In *Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra*, vol. 1, no. 1, pp. 11-18, 2021
- [7] Y.S. Mahardhika, E. Zuliarso, "Analisis Sentimen Terhadap Pemerintahan Joko Widodo Pada Media Sosial Twitter Menggunakan Algoritma Naives Bayes Classifier". *Prosiding SINTAK 2018*, pp. 409-413, 2018

- 
- [8] P.P.E. Indarbensyah, N. Rochmawati, "Penerapan N-Gram Menggunakan Algoritma Random Forest Dan Naïve Bayes Classifier Pada Analisis Sentimen Kebijakan PPKM 2021". *J. Informatics Comput. Sci*, vol. 2, no. 4, pp. 235-244, 2021.
- [9] M.S. Mustafa, M.R. Ramadhan, & A.P. Thenata, "Implementasi Data Mining untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes Classifier". *Creative Information Technology Journal*, vol. 4, no. 2, pp. 151-162, 2018.
- [10] D. Darwis, N. Siskawati, & Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional". *Jurnal Tekno Kompak*, vol. 15, no. 1, pp. 131-145, 2021.
- [11] A. Damuri, U. Riyanto, H. Rusdianto, & M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako". *JURIKOM (Jurnal Riset Komputer)*, vol. 8, no. 6., pp. 219-225, 2021.
- [12] W.E. Nurjanah, R.S. Perdana, & M.A. Fauzi, "Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet". *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, pp. 1750-1757, 2017
- [13] F.F. Rachman, & S. Pramana, "Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter". *Indonesian of Health Information Management Journal (INOHIM)*, vol. 8, no. 2, pp. 100-109, 2020.
- [14] Sya'bani, M.R. Fais, U. Enri, & T.N. Padilah. "Analisis Sentimen Terhadap Bakal Calon Presiden 2024 Dengan Algoritme Naïve Bayes." *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 2, pp. 265-273, 2022.
- [15] Mimboro, Prasetyo. "Sentimen Twitter terhadap PILKADA kota Medan menggunakan metode Naive Bayes." *JNANALOKA*, vol. 3, no. 1, pp. 27-32, 2022.
- [16] Nurhazizah, Eva, R.N. Ichsan, & S. Widiyanesti, "Analisis Sentimen Dan Jaringan Sosial Pada Penyebaran Informasi Vaksinasi Di Twitter", *J Swabumi*, vol. 10, no. 1, pp. 24-35, 2022
- [17] A. Nugroho, "Analisis Sentimen Pada Media Sosial Twitter Menggunakan Naive Bayes Classifier Dengan Ekstrasi Fitur N-Gram". *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, vol. 2, no. 2, pp. 200-209, 2018.
- [18] M. Muljono, D.P. Artanti, A. Syukur, & A. Prihandono, "Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naïve Bayes". *Konferensi Nasional Sistem Informasi (KNSI) 2018*, pp. 165-170, 2018.