

Analisis Sentimen Komentar Konten Kreator Gaming Menggunakan Metode *Naive Bayes* dan KNN

M. Dimas Taufiqurahman^a, Siska Anraeni^b, Herdianti Darwis^c

Universitas Muslim Indonesia, Makassar, Indonesia

Email:^a13020200274@umi.ac.id; ^bsiskaanraeni@umi.ac.id; ^cherdianti.darwis@umi.ac.id

Received: 10-07-2024 | Revised: 20-08-2024 | Accepted: 10-11-2024 | Published: 29-12-2024

Abstrak

Perkembangan teknologi telah membawa banyak pergeseran termasuk dalam transformasi media konvensional seperti radio dan televisi menjadi media online. Salah satu platform yang mengalami peningkatan signifikan adalah *YouTube* yang kini menjadi pilihan utama hiburan, terutama bagi anak-anak dan orang dewasa. *YouTube*, didirikan pada Februari 2005, memiliki 2.56 miliar pengguna per Januari 2022 dan memuat berbagai konten, termasuk video klip film, acara TV, dan video buatan pengguna. Di Indonesia, konten tutorial game menjadi salah satu yang paling sering ditonton, mencapai 7.1% dari total 19 jenis konten populer berdasarkan survei dari Asosiasi Penyelenggara Jasa Internet Indonesia (APJII). Penelitian ini bertujuan untuk mengekstrak data opini dan menganalisis sentimen *netizen* terhadap konten *gaming* di *YouTube* menggunakan pendekatan *Naive Bayes* dan KNN. Berdasarkan hasil klasifikasi kedua metode *Naive Bayes* & KNN didapatkan hasil akurasi 96% untuk metode *Naive Bayes* dan akurasi 87% untuk metode KNN sehingga dapat disimpulkan bahwa metode *Naive Bayes* memiliki akurasi yang lebih tinggi.

Kata kunci: *Youtube*, *Gaming*, *Naive Bayes*, *K-Nearest Neighbor*, *Sentiment*

Pendahuluan

Perkembangan media baru yang pesat telah mengakibatkan perubahan besar pada media konvensional, termasuk peralihan radio dan televisi ke format online. Seiring waktu, televisi semakin kurang diminati, dan kegiatan menonton masyarakat beralih ke *Youtube* [1]. Hadirnya *Youtube* menawarkan hiburan baru yang menarik terkhususnya bagi anak-anak dan orang dewasa. Kini anak-anak lebih menyukai menonton konten favorit mereka melalui *Youtube* membuat *Youtube* banyak dilihat oleh anak-anak salah satunya konten *gaming* [2].

Youtube merupakan salah satu *platform* media sosial yang populer dan banyak diminati pada masa ini. Dalam *platform* ini, berbagai konten video diunggah dan berhasil menarik perhatian sejumlah besar pengguna [3]. Selain itu, jumlah pengguna *Youtube* mencapai angka yang sangat signifikan. Berdasarkan laporan *We Are Social* pada bulan Januari 2022, *platform* ini memiliki 2,56 miliar pengguna. Penting untuk dicatat bahwa *Youtube* didirikan pada bulan Februari 2005 oleh tiga mantan karyawan *Paypal*. *Youtube* umumnya memuat klip video yang berasal dari film, acara TV, dan karya video yang dibuat oleh pengguna [4]. Data yang diperoleh dari *goodstats.id* pada tahun 2022 juga mendukung fakta-fakta tersebut. Konten Tutorial Game menjadi salah satu Konten *Youtube* yang paling sering ditonton *Netizen* Indonesia. Berdasarkan survei yang disebar oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII), data yang terkumpul mengenai konten *Youtube* yang paling sering ditonton oleh pengguna internet Indonesia ada 19 macam. Dari 19 konten, terdapat 10 konten yang paling sering ditonton dan Tutorial Game menjadi salah satunya dengan mencapai persentase sebanyak 7,1%. Atas dasar berbagai respon dari *netizen* yang disampaikan melalui kolom komentar pada konten kreator *gaming* membuat penulis merasa tertarik untuk menyelidiki respon negatif *netizen* dan juga isi kontennya yang problematik dengan *netizen* yang menyaksikan kontennya [5].

Salah satu metode untuk memahami tanggapan komentar masyarakat di media *Youtube* adalah melalui analisis sentimen. Analisis sentimen adalah cabang dari text mining yang digunakan untuk menilai dan mengekstrak opini dari data teks. Teknik ini secara otomatis memproses teks untuk menentukan makna sentimen yang terdapat dalam opini atau tulisan, yang dikategorikan ke dalam sentimen positif, netral atau negatif [6].

Penelitian [7] Analisis sentimen diterapkan menggunakan metode *Naive Bayes Classifier* dan *TextBlob* dari pustaka *TextBlob* dalam bahasa pemrograman Python. Dataset yang diperoleh dari Twitter berjumlah 405 data, dengan hasil sentimen negatif mencapai 70,62% atau 286 data, sentimen positif sebesar 22,72% atau 92 data, dan sentimen netral sebesar 6,67% atau 27 data. Sementara itu, metode *TextBlob* menunjukkan hasil sentimen negatif yang lebih tinggi yaitu 55,8% atau 226 data, sentimen positif 33,09% atau 134 data, dan sentimen

netral 11,11% atau 45 data. Ini mengindikasikan bahwa masyarakat merasa kurang puas terhadap penggunaan paylater. Pengujian model menggunakan confusion matrix menunjukkan bahwa algoritma Naive Bayes Classifier memiliki akurasi lebih tinggi, yaitu 91%, dibandingkan dengan TextBlob yang hanya 61%. Berdasarkan temuan tersebut, penelitian ini bertujuan untuk mengekstrak data opini dan menganalisis sentimen netizen terhadap konten gaming di YouTube menggunakan pendekatan Naive Bayes dan K-Nearest Neighbors (KNN).

Metode

Naive bayes adalah metode klasifikasi yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. Metode tersebut berlandaskan pada *teorema Bayes* dan memiliki kemampuanklasifikasi yang sebanding dengan *decision tree* dan *neural network* [8].

$$P(C_i|X) = \frac{p(X|C_i).P(C_i)}{P(X)} \quad (1)$$

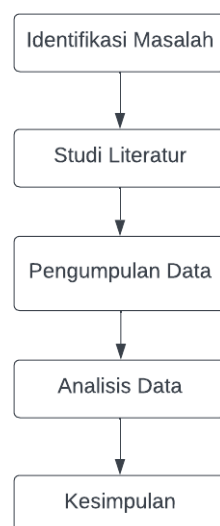
Naive Bayes adalah metode klasifikasi sederhana yang sering digunakan karena kemudahannya dalam *penerapan* dan kemampuannya dalam memberikan hasil yang baik dalam berbagai kasus. Namun, kelemahan dari metode ini adalah asumsi atau dengan kata lain kondisi kelas saling bebas, sehingga kurang akurat. Adapun pada prakteknya, ketergantungan ada diantara variabel [9].

Algoritma *K-Nearest Neighbors* (KNN) merupakan metode yang digunakan untuk *mengklasifikasikandata* berdasarkan jarak terdekat antara objek data. Penentuan nilai K yang optimal untuk algoritma ini berdasarkan pada data yang tersedia. Nilai K yang tinggi dapat mengurangi pengaruh efek noise pada klasifikasi, bisa juga membuat batasan antara setiap kategori [10].

$$D_{xy} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2)$$

Jarak kedekatan dilambangkan sebagai *D*, merupakan ukuran yang digunakan untuk menentukan seberapa mirip dua data, yaitu data *training* (*x*) dan data *testing* (*y*). Data ini memiliki sejumlah atribut individu (*n*). Fungsi similaritas atribut (*f*) digunakan untuk menghitung tingkat kemiripan atribut ke-*i* antara kasus *X* dan kasus *Y* [11]. Atribut individu ini juga diberi indeks *i*, yang bernilai mulai dari 1 sampai dengan *i*. Penggunaan jarak kedekatan dan fungsi similaritas ini sangat penting dalam berbagai algoritma *machine learning* dan *data mining* untuk mengevaluasi performa model atau untuk melakukan prediksi berdasarkan kemiripan data.

Penulis memberikan gambaran terhadap alur penelitian yang dilakukan. Tahapan penelitian yang dibangun ditujukan pada Gambar 1.

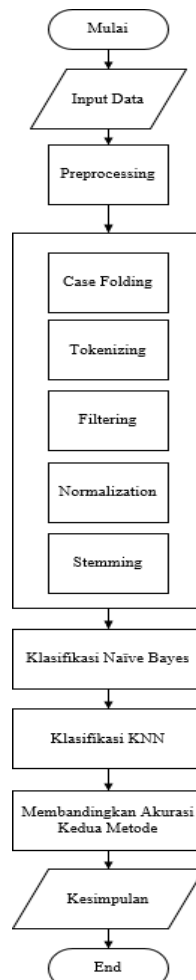


Gambar 1. Tahapan Penelitian

Penelitian ini dimulai dengan mengidentifikasi masalah pada komentar YouTube konten kreator gaming. Tahap selanjutnya adalah studi literatur dengan memahami penelitian serupa yang pernah dilakukan sebelumnya untuk mendapatkan teori dan informasi yang mendukung mengenai klasifikasi komentar negatif, positif, dan

netral. Pengumpulan data dilakukan menggunakan *YouTube* API untuk memperoleh data yang diperlukan selama penelitian. Selanjutnya, dilakukan analisis data dengan menggunakan klasifikasi *Naive Bayes* dan KNN untuk membandingkan akurasi kedua metode. Library yang digunakan pada *google colab* meliputi *stopwords* dan *stemming* list bahasa Indonesia, serta nilai akurasi dari metode tersebut.

Perancangan



Gambar 2. Desain Penelitian

A. Input Data

Pada penelitian ini yang dilakukan pertama kali penginputan data yang diambil dari kolom komentar konten kreator gaming menggunakan *Youtube* API.

B. Preprocessing

Pada tahap preprocessing ini, data yang telah di crawling melalui Youtube API akan diolah menjadi data yang dapat digunakan pada tahapan selanjutnya. Tahapan *Preprocessing* terdiri dari *case folding*, *tokenizing*, *filtering* dan *stemming*.

C. Klasifikasi *Naive Bayes* dan KNN

Tahapan selanjutnya yaitu melakukan analisis dengan metode *Naive Bayes* dan KNN dengan mendapatkan tingkat akurasi dalam menentukan sentimen.

D. Evaluasi

Setelah melakukan klasifikasi tahapan selanjutnya yaitu evaluasi, Ada beberapa hal yang bisadigunakan untuk mengevaluasi kinerja klasifikasi. Penelitian ini menggunakan *Accuracy Score* sebagai parameter evaluasi untuk mengetahui seberapa baik hasil yang akan diberikan.

Tahap *preprocessing* terdiri dari empat langkah utama, pertama, data *cleaning* dan *case folding*, yang berfungsi untuk membersihkan data dari kesalahan termasuk noise data dan mengubah semua teks menjadi huruf kecil. Kedua, *filtering*, yang menghapus kata-kata atau karakter yang tidak relevan dan rutin muncul, juga menghilangkan slang dan kata-kata pendek. Ketiga, *stemming*, yang mengambil kata dasar atau akar kata untuk

membantu dalam klasifikasi teks. Keempat, *tokenizing*, yang mengubah kalimat menjadi makna yang lebih spesifik menggunakan konsep *bigram*, yaitu dua kata yang berpasangan [12]. Yang dimana tahapan *preprocessing* dalam hal ini meliputi :

1. Proses *Case Folding*

Tabel 1. Penerapan *Case Folding*

Sebelum	Sesudah
Menurut saya, Gameplay yang diperlihatkan cukup baik, Sangat membantu.	menurut saya, gameplay yang diperlihatkan cukup baik, sangat membantu.

2. Proses *Tokenizing*

Tabel 2. Penerapan *Tokenizing*

Sebelum	Sesudah
menurut saya, gameplay yang diperlihatkancukup baik, sangat membantu.	menurut, saya, gameplay, yang, diperlihatkan, cukup,baik, sangat, membantu

3. Proses *Filtering*

Tabel 3. Penerapan *Filtering*

Sebelum	Sesudah
menurut, saya, gameplay, yang, diperlihatkan, cukup, baik, sangat, membantu	Gameplay di perlihatkan cukup baik membantu

4. Proses *Normalization*

Tabel 4. Penerapan *Normalization*

Sebelum	Sesudah
Gameplay di perlihatkan cukup baik membantu	Gameplay ditampilkan dengan cukup baik dan membantu.

5. Proses *Stemming*

Tabel 5. Penerapan *Stemming*

Sebelum	Sesudah
menurut saya, gameplay yang diperlihatkan cukup baik, sangat membantu.	menurut saya gameplay yang lihat cukup baik sangatbantu

Pemodelan

A. *Crawling* Data

Melakukan *crawling* data menggunakan dataset dari kolom komentar dari salah satu konten kreator *gaming* dengan cara menggunakan Youtube API agar bisa mendapatkan akses untuk melakukan proses *crawling* data dari *platform Youtube*. Setelah proses *crawling* data dilakukan, akan dilanjutkan ke proses yang selanjutnya yaitu *Text preprocessing* untuk dilakukan proses pembersihan data yang telah di *crawling* sebelumnya.

B. *Preprocessing*

Preprocessing adalah langkah awal yang mengubah teks mentah menjadi input dan menerapkan berbagai rutinitas dasar untuk mengubah atau menghapus elemen-elemen tekstual yang tidak relevan untuk pengolahan selanjutnya [13]. Berikut adalah beberapa proses dari *preprocessing*:

1. *Case folding*

Pada Gambar 3 adalah proses *Case Folding* proses ini melibatkan penyamaan format huruf dalam sebuah dokumen. Tahapan ini dilakukan karena tidak semua dokumen teks konsisten dalam penggunaan huruf kapital [14]. Oleh karena itu, diperlukan peran *case folding* dalam mengubah keseluruhan teks dalam suatu dokumen dari huruf kapital menjadi bentuk standar yang umumnya berupa huruf kecil

Case Folding Result :

	textDisplay
0	munafikla adi katanya gak baperan wkwkw
1	udah si pling bijak dan bener aja si adi ni
2	adi itu kan cuma numpang di onic doank ges.. o...
3	kent kash snper kepala s ad
4	yang setuju tahun depan bken kurban loke
...	...
295	real gue nonton react ini krena tumbnail
296	guwe gak gitu omgnya
297	siapa pun coach nya yg melatih onic jika masih...
298	the real apa kawan
299	rrq semenjak di dukung brandon kent lebih bany...
300 rows × 1 columns	

Gambar 3. Proses *Case Folding*

2. Proses *Tokenizing*

Pada Gambar 4 adalah proses *Tokenizing* yang merupakan proses memisahkan setiap kata yang membentuk dokumen [15] percakapan. Proses ini menghilangkan tanda baca seperti titik (.), Koma (,), tanda tanya (?), dan lain-lain.

```
Tokenizing Result :
0      [munafikla, adi, katanya, gak, baperan, wkwkw]
1      [udah, si, pling, bijak, dan, bener, aja, si, ...
2      [adi, itu, kan, cuma, numpang, di, onic, doank...
3      [kent, kash, snper, kepala, ad]
4      [yang, setuju, tahun, depan, bken, kurban, loke]
5      [ngeliat, si, adi, jadi, keinget, ernada, espr...
6      [adi, emng, anj, si]
7      [malu, banget, bayangin, jd, si, acil]
8      [hantam, aja, bang, bikent]
9      [sebelum, adi, masuk, onic, kan, onic, emang, ...
10     [gw, yakin, kalo, adi, ga, di, onic, onic, tet...
11     [kocak, acl, bn, ad, emng]
12     [tu, caca, bukan, vonsey]
13     [yutuber, plak, ketiplek, diam, aja]
14     [kok, marah, marah, wkwk]
15     [yaelah, klo, gasuka, jangan, di, tonton, lah,...
16     [ngejilat, ludah, sendiri]
17     [gas, adu, tinju, di, byon, combat, 4, ken]
18     [tenang, kent, si, adi, juga, ngk, ad, ngaruhn...
19     [gassss, tinju]
Name: textDisplay_tokens, dtype: object
```

Gambar 4. Hasil *Tokenizing*

3. Proses *Filtering*

Pada Gambar 5 adalah proses *Filtering* atau proses menghilangkan kata-kata yang tidak dibutuhkan atau kata-kata yang tidak memiliki arti [16].

```
0      [munafikla, adi, baperan, wkwkw]
1      [udah, pling, bijak, bener, adi, ni]
2      [adi, numpang, onic, doank, ges, onic, adi, jg...
3      [kent, kash, snper, kepala, ad]
4      [setuju, bken, kurban, loke]
5      [ngeliat, adi, keinget, ernada, espro, wkwkwkwk]
6      [adi, emng, anj]
7      [malu, banget, bayangin, acil]
8      [hantam, bang, bikent]
9      [adi, masuk, onic, onic, emang, udah, kuatgaus...
10     [gw, adi, onic, onic, tetep, strong]
11     [kocak, acl, bn, ad, emng]
12     [tu, caca, vonsey]
13     [yutuber, plak, ketiplek, diam]
14     [marah, marah, wkwk]
15     [yaelah, gasuka, tonton, tinggal, skip, susahn...
16     [ngejilat, ludah]
17     [gas, adu, tinju, byon, combat, 4, ken]
18     [tenang, kent, adi, ngk, ad, ngaruhnya, onic]
19     [gassss, tinju]
Name: textDisplay_tokens_WSW, dtype: object
```

Gambar 5. Proses *Filtering*4. Proses *Normalization*

Pada Gambar 6 adalah proses *Normalization* yang merupakan proses mengubah kata tidak standar seperti kata gaul, atau ditingkatkan menjadi kata formal [17].

```

0      [munafikla, adik, baperan, wkwwk]
1      [sudah, pling, bijak, benar, adik, ini]
2      [adik, numpang, onic, doank, teman, onic, adik...]
3      [kent, kash, snper, kepala, ada]
4      [setuju, bken, kurban, loke]
5      [ngeliat, adik, keinget, ernada, espro, wkwwkwwk]
6      [adik, memang, anjing]
7      [malu, sekali, bayangin, acil]
8      [hantam, kakak, bikent]
9      [adik, masuk, onic, onic, memang, sudah, kuatg...]
10     [aku, adik, onic, onic, tetap, strong]
11     [kocak, acl, bn, ada, memang]
12     [itu, caca, vonsey]
13     [yutuber, plak, ketiplek, diam]
14     [marah, marah, wkwwk]
15     [yaelah, gasuka, tonton, tinggal, skip, susahn...]
16     [ngejilat, ludah]
17     [gas, adu, tinju, byon, combat, 4, ken]
18     [tenang, kent, adik, tidak, ada, ngaruhnya, onic]
19     [gassss, tinju]
Name: textDisplay_normalized, dtype: object

```

Gambar 6. Proses *Normalization*5. Proses *Stemming*

Stemming merupakan proses mengubah kata berimbuhan menjadi bentuk dasar. Proses ini dapat mengakibatkan hilangnya makna, karena sebuah kata dapat memiliki arti yang berbeda setelah ditambahkan imbuhan. Gambar 7 menunjukkan hasil dari proses *stemming*.

```

0      munafikla adi baperan wkwwk
1      udah pling bijak bener adi ni
2      adi numpang onic doank ges onic adi jg gk bkla...
3      kent kash snper kepala ad
4      tuju bken kurban loke
...
295     real gue nonton react krena thumbnail
296     guwe gitu orngnya
297     coach latih onic onic kiboy sanz percaya dah o...
298     the real kawan
299     rrq semenjak dukung brandon kent drama prestasi
Name: Reviews_clean, Length: 300, dtype: object

```

Gambar 7. Proses *Stemming*C. *Wordcloud*1. *Wordcloud*

Wordcloud digunakan untuk menampilkan citraan visual dari teks tertulis dan dapat digunakan sebagai sebuah alat bantu dalam analisis teks [18]. Kata-kata yang ada digambarkan ke dalam suatu visualisasi dimana besar kecilnya huruf bergantung pada frekuensi munculnya kata tersebut di dalam teks.

a. *Wordcloud* Sentimen Positif

Visualisasi *wordcloud* sentimen positif pada Sentimen komentar konten kreator *gaming* menciptakan gambaran grafis yang menyoroti kata-kata kunci yang sering muncul dalam konteks positif [19].

b. *Wordcloud* Sentimen Netral

Wordcloud sentimen netral memberikan gambaran visual atas kata-kata yang paling umum dalam ulasan dengan sentimen netral terhadap sentimen komentar konten kreator *gaming*.

c. *Wordcloud* Sentimen Negatif

Melalui *wordcloud* sentimen negatif, kata-kata yang dominan dalam ulasan dengan sentimen negatif yang dipresentasikan secara visual. Ini memberikan wawasan terhadap masalah atau ketidakpuasan konten kreator *gaming*.



Gambar 8. *Wordcloud* sentimen positif

Pada Gambar 8 merupakan *wordcloud* komentar kreator gaming yang sering muncul dalam sentimen positif yaitu kata adi, onic, aja, bang, coach, baper dan acil.



Gambar 9. *Wordcloud* sentimen netral

Pada Gambar 9 merupakan *wordcloud* komentar kreator gaming yang sering muncul dalam sentimen netral yaitu kata adi, onic, rrq, juara, tumpengan, acil, coach.



Gambar 10. *Wordcloud* sentimen negatif

Pada Gambar 10 merupakan *wordclod* komentar kreator gaming yang sering muncul dalam sentimen negatif yaitu kata adi, onic, aja, rrq, gw, jago dan numpang.

2. Menghitung TF-IDF

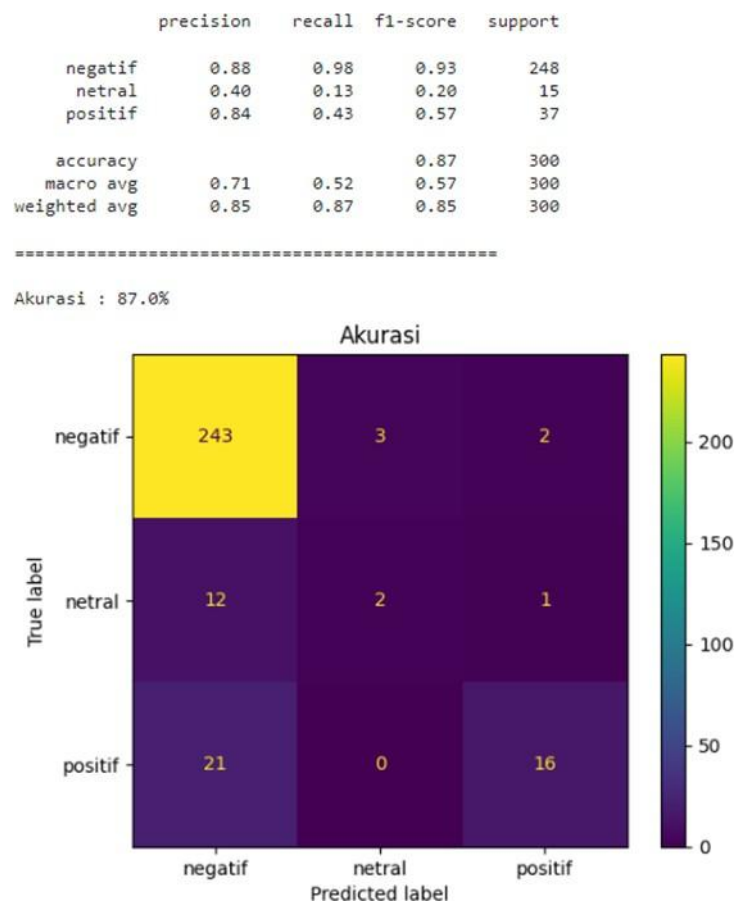
IDF mengukur pentingnya suatu term dalam seluruh korpus. Biasanya, IDF dihitung dengan mengambil log dari perbandingan antara jumlah total dokumen dalam korpus dan jumlah dokumen yang mengandung term tersebut [20]. Daftar istilah dengan nilai TF-IDF dapat dilihat pada Gambar 11.

wkwkw	-	0.45733997717709995
baperan	-	0.43388119855036705
gak	-	0.29866246548803654
katanya	-	0.4904032686464286
adi	-	0.17999236297840987
munafikla	-	0.4904032686464286
ni	-	0.0
aja	-	0.0
bener	-	0.0
dan	-	0.0
bijak	-	0.0
pling	-	0.0
si	-	0.0
udah	-	0.0
adi	-	0.17999236297840987
prformanya	-	0.0
turun	-	0.0
bklan	-	0.0
gk	-	0.0
jg	-	0.0
tanpa	-	0.0
ges	-	0.0
doank	-	0.0
onic	-	0.0

Gambar 11. Hasil Proses TF-IDF

D. Evaluasi Model

Pada tahap ini dilakukan dalam menguji performa dari tiap metode KNN dan *naïve bayes* dengan membagi data training dan data testing. Dalam pengukuran model penulis menggunakan *Confusion Matrix* sehingga menghasilkan nilai akurasi pada Gambar 12.



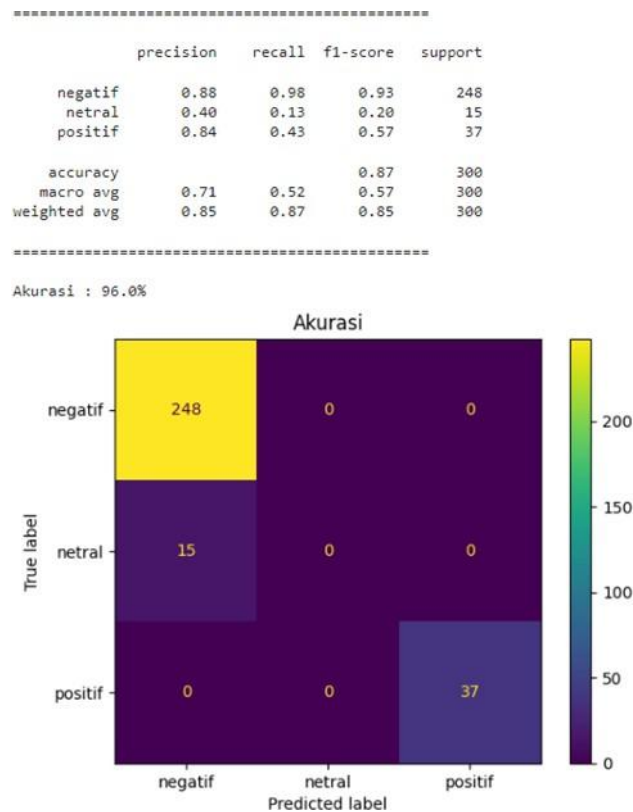
Gambar 12. Hasil Akurasi KNN

Pada Gambar 12 merupakan hasil dari perhitungan akurasi dari metode KNN yang mana pada hasilakhirnya mendapatkan hasil 87% akurasi.

Tabel 6. *Confusion Matrix* KNN

	Negatif	Netral	Positif
Negatif	243	3	2
Netral	12	2	1
Positif	21	0	16

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} = \frac{16 + 243}{16 + 243 + 2 + 1} = \frac{259}{296} = 87\%$$



Gambar 13. Hasil Akurasi *Naive Bayes*

Pada Gambar 12 merupakan hasil dari perhitungan akurasi dari metode *Naive Bayes* yang mana pada hasilakhirnya mendapatkan hasil 96% akurasi.

Tabel 7. *Confusion Matrix* *Naive Bayes*

	Negatif	Netral	Positif
Negatif	248	0	0
Netral	15	0	0
Positif	0	0	37

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} = \frac{248 + 37}{248 + 15 + 37 + 0} = \frac{285}{300} = 96\%$$

Kesimpulan

Berdasarkan hasil klasifikasi kedua metode *Naïve Bayes* & KNN didapatkan hasil akurasi 96% untuk metode *Naïve Bayes* dan akurasi 87% untuk metode KNN sehingga dapat disimpulkan bahwa metode *Naïve Bayes* memiliki akurasi lebih tinggi.

Daftar Pustaka

- [1] Dini Wahdiyati And Reyvianto Dwi Putra, “Kekerasan Verbal Dalam Konten Gaming Di Youtube (Analisis Isi Kualitatif Konten Ulasan Permainan Online Maincraft Dan Mobile Legend Pada Akun Youtube Miuveox Dan Brandonkent Everything),” *Jurnal Indonesia Sosial Teknologi*, Vol. 3, Pp. 203–218, 2022.
- [2] M. Jamil Reza, “Persepsi Mahasiswa Terhadap Penggunaan Youtube Sebagai Media Konten Video Kreatif,” *Jurnal Komunikasi Dan Organisasi (J-Ko)*, Vol. 3, Pp. 39–46, 2021.
- [3] A. Felix, J. Wicaksana Warda, F. Jonathan, F. Nicolass Hadi, And S. Surianto Putra, “Analisis Faktor Dan Strategi Branding Youtuber Jerome Polin Dalam Mempertahankan Eksistensi,” *Jurnal Ilmu Politik Dan Komunikasi*, Vol. 13, No. 2, Pp. 115–127, 2023.
- [4] D. Kanal, Y. Gerrard, W. Ferdianto, T. Nugroho, And S. N. Febriyanti, “Analisis Wacana Kritis Trash-Talk Pada Konten Game Mobile Legends,” *Jurnal Ilmiah Wahana Pendidikan*, Oktober, Vol. 2023, No. 19, Pp. 277–294, Doi: 10.5281/Zenodo.8388761.
- [5] F. Ushuludin Adab Dan Dakwah, “Skripsi Analisis Respon Netizen Terhadap Konten Akun Youtube Zavildatv Diajukan Untuk Memenuhi Syarat-Syarat Guna Memperoleh Gelar Sarjana (S-1) Pada Ilmu Dakwah.”
- [6] Habibi Al Rasyid Harpizon, Rahmad Kurniawan, Iwan Iskandar, Roni Salambue, Elvia Budianita, And Fadhilah Syafria, “Analisis Sentimen Komentar Di Youtube Tentang Ceramah Ustadz Abdul Somad Menggunakan Algoritma Naive Bayes,” *Jurnal Nasional Komputasi Dan Teknologi Informasi*, Vol. 5, Pp. 131–140, 2022.
- [7] A. Safira, A. S. Masyarakat..., And F. N. Hasan, “Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier,” *Jurnal Sistem Informasi*, Vol. 5, No.1, 2023.
- [8] H. F. Putro, R. T. Vlandari, And W. L. Y. Saptomo, “Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan,” *Jurnal Teknologi Informasi Dan Komunikasi (Tikomsin)*, Vol. 8, No.2, Oct. 2020, Doi: 10.30646/Tikomsin.V8i2.500.
- [9] N. Chatrina Siregar, R. Ruli, A. Siregar, ; M Yoga, And D. Sudirman, “Implementasi Metode Naive Bayes Classifier (Nbc) Pada Komentar Warga Sekolah Mengenai Pelaksanaan Pembelajaran Jarak Jauh (Pjj),” 2020.
- [10] S. R. Cholil, T. Handayani, R. Prathivi, And T. Ardianita, “Implementasi Algoritma K-Nearest Neighbor (Knn) Untuk Klasifikasi Seleksi Penerima Beasiswa,” *Ijcit (Indonesian Journal On Computer And Information Technology)*, Vol. 6, No. 2, Pp. 118–127, 2021.
- [11] Neutrino Sae B. Kusrorong, Derwin R. Sina2, And Nelci Dessy Rumlaklak, “Kajian Machine Learning Dengan Komparasi Klasifikasi Prediksi Dataset Tenaga Kerja Non-Aktif,” *J-Icon*, Vol. 7, No. 1, Pp. 37–49, 2019.
- [12] Herdianti Darwis, Nugraha Wanaspati, And Siska Anraeni, “Support Vector Machine Untuk Analisis Sentimen Masyarakat Terhadap Penggunaan Antibiotik Di Indonesia,” *Indonesian Journal Of Computer Science*, Vol. 12, Pp. 2196–2206, 2023.
- [13] Hadiyatun Najjichah, Abdul Syukur, And Hendro Subagyo, “Pengaruh Text Preprocessing Dan Kombinasinya Pada Peringkat Dokumen Otomatis Teks Berbahasa Indonesia,” *Teknologi Informasi*, Vol. 15, Pp. 1–11, 2019.
- [14] D. Alita And A. Rahman, “Pendeteksian Sarkasme Pada Proses Analisis Sentimen Menggunakan Random Forest Classifier,” 2020.
- [15] H. Prasetyo And A. Senja Fitriani, “Sentiment Analysis Before Presidential Election 2024 Using Naïve Bayes Classifier Based On Public Opinion In Twitter [Analisa Sentimen Jelang Pilpres 2024 Menggunakan Naïve Bayes Classifier Berdasarkan Opini Publik Di Twitter],” 2024.
- [16] A. H. Tri Jaka, “Preprocessing Text Untuk Meminimalisir Kata Yang Tidak Berarti Dalam Proses

- Text Mining,” *Jurnal Informatika Upgris*, Vol. 1, No. 6, Pp. 1–9, 2018.
- [17] T. Meisya, P. Aulia, N. Arifin, And R. Mayasari, “Perbandingan Kernel Support Vector Machine (Svm) Dalam Penerapan Analisis Sentimen Vaksinisasi Covid-19,” *Sintech Journal* Vol. 4, No. 2, Pp. 139–145, 2021, [Online]. Available: <https://doi.org/10.31598>
- [18] S. Febriyanti, D. R. Nursidah, D. Gustiara, And R. Yulianti, “Analisis Sentimen Masyarakat Indonesia Terhadap Pelaksanaan Vaksin Covid’19,” *Jurnal Khazanah*, Vol. 13, No. 2, 2021, [Online]. Available: <https://journal.uii.ac.id/khazanah>
- [19] I. Tri Julianto, “Analisis Sentimen Terhadap Sistem Informasi Akademik Mahasiswa Institut Teknologi Garut,” *Jurnal Algoritma*, Vol. 19, No. 1, Pp. 458–465, 2022, [Online]. Available: <https://jurnal.itg.ac.id/>
- [20] D. Septiani And I. Isabela, “Analisis Term Frequency Inverse Document Frequency (Tf-Idf) Dalam Temu Kembali Informasi Pada Dokumen Teks,” *Sintesia: Jurnal Sistem Dan Teknologi Informasi Indonesia*, Vol. 1, No. 2, Pp. 81–88, 2022.