

ANALISIS SENTIMEN KOMENTAR NETIZEN TWITTER TERHADAP FANBOY NCT MENGGUNAKAN NAÏVE BAYES CLASSIFIER

Efrina Ashfadiniyya Ardisavira, Dewi Sri Andika Rusmana, Herlina Kusumaningrum

Program Studi Ilmu Komunikasi, Universitas 17 Agustus 1945 Surabaya

e-mail: efrinaqc@gmail.com

Abstract

Twitter is one of the most widely accessed social media in Indonesia. Users can share messages and information discussing various topics or issues. This research uses a quantitative content approach and Naïve Bayes Classifier sentiment analysis method to find out how Twitter netizens comment on NCT fanboys. Twitter data was collected through crawling data techniques with the tweet keyword "fanboy nct" in Indonesian language between October 3, 2022 – April 3, 2023. The labeling of positive, negative, and neutral sentiments in the text is determined based on the concept of attitude change in the Social Judgment Theory proposed by Muzafer Sherif. From 2 (two) data scenarios involving 408 clean datasets, scenario 1 became the most optimal experiment that resulted in sentiment positive 70,3% (287 data), negative 6,1% (25 data), and neutral 23,5% (96 data). The evaluation of confusion matrix obtained precision class 95,07%, recall 98,72%, and accuration 98,28%. The high test percentage value indicates that the performance of Naïve Bayes is stable and excellent in classifying sentiment.

15

Keywords: sentiment analysis, NCT fanboys, Naïve Bayes Classifier, RapidMiner, Twitter

Abstrak

Twitter merupakan salah satu media sosial yang banyak diakses di Indonesia. Para penggunanya dapat saling bertukar pesan dan informasi mendiskusikan bermacam topik atau isu. Penelitian ini menggunakan pendekatan kuantitatif isi metode analisis sentimen Naïve Bayes Classifier untuk mengetahui bagaimana komentar netizen Twitter terhadap fanboy NCT. Data Twitter dikumpulkan lewat teknik *crawling data* dengan kata kunci *twit "fanboy nct"* berbahasa Indonesia rentang waktu 3 Oktober 2022 – 3 April 2023. Pelabelan sentimen positif, negatif, dan netral pada teks ditentukan berdasarkan konsep perubahan sikap Teori Penilaian Sosial yang dikemukakan oleh Muzafer Sherif. Dari 2 (dua) skenario data yang melibatkan 408 dataset bersih, skenario 1 adalah percobaan paling optimal yang menghasilkan sentimen positif 70,3% (287 data), negatif 6,1% (25 data), dan netral 23,5% (96 data). Evaluasi *confusion matrix* memperoleh *precision* 95,07%, *recall* 98,72%, dan *accuration* sebesar 98,28%. Tingginya nilai persentase pengujian memperlihatkan bahwa kinerja Naïve Bayes sangat baik dan stabil dalam melakukan klasifikasi sentimen.

3

Kata kunci: analisis sentimen, fanboy NCT, klasifikasi Naïve Bayes, RapidMiner, Twitter

PENDAHULUAN

Latar Belakang

Kemajuan teknologi informasi dan komunikasi membawa pengaruh besar di berbagai aspek kehidupan, tidak terkecuali budaya populer. Salah satu budaya populer yang beberapa tahun terakhir menerima sorotan adalah *K-Pop*. *K-Pop* singkatan dari *Korean Pop* atau *Korean Popular Music*, yaitu musik pop asal Korea Selatan. Dalam perkembangannya, *K-Pop* menjadi bagian yang tak bisa dipisahkan dari *Hallyu Wave* atau *Korean Wave* (Gelombang Korea). Istilah *Hallyu Wave* dicetuskan media China tahun 1997 untuk menyebut produk-produk kebudayaan Korea (Ri'aeni et al., 2019).

K-Pop diperkirakan menyebar di Indonesia pertengahan 2000-an. Demam *K-Pop* mulai mendunia setelah lagu "Gangnam Style" dari PSY yang rilis pada 15 Juli 2012 booming hingga memuncaki tangga lagu di pelbagai negara (Rizal, 2022). Meningkatnya tren *K-Pop* tentu tidak lepas dari pro dan kontra. Di Indonesia, sebagian masyarakat tidak memperlakukan masuknya budaya Korea Selatan tersebut. Penggemar *K-Pop* dilihat sebagai komunitas yang solid dan loyal terhadap sang idola, sering melakukan penggalangan dana membantu sesama yang membutuhkan, dan giat bersuara menanggapi isu politik, hukum, kesehatan, dan sebagainya. Pun sebaliknya, terdapat masyarakat yang tidak menyukai *K-Pop* karena dikhawatirkan akan menggeser budaya asli Indonesia. Pihak kontra menilai penggemar *K-Pop* fanatik, kurang peduli musik lokal, boros membeli *merchandise K-Pop*, dan menimbulkan

keributan antarpenggemar di media sosial (Haloho et al., 2023). Perbedaan pendapat atau opini merupakan hal lumrah sebab tingkat penerimaan ataupun penolakan setiap orang mengenai suatu isu didasari bermacam faktor, seperti kepercayaan, pengalaman, keterlibatan ego, dan faktor lain (Morissan, 2013).

Korea Foundation dalam Analisis Status Global Hallyu 2022 menyebutkan jumlah penggemar *Hallyu* di seluruh dunia mencapai 178.825.261 orang. Angka ini naik 19 kali lipat dibandingkan 11 tahun lalu. *K-Pop* memangku kekuatan pendorong dibalik pertumbuhan *Hallyu* (Won, 2023). Sampai saat ini, mayoritas penggemar *K-Pop* masih dipegang kaum perempuan. Hasil survei IDN Times yang melibatkan 580 responden penggemar *K-Pop* di Indonesia sebanyak 92,1% penggemar *K-Pop* adalah perempuan, sedangkan laki-laki 7,9% (Triadanti, 2019). Selisih itu membuktikan *fanboy K-Pop* tergolong minoritas.

Fanboy didefinisikan sebagai laki-laki yang menggemari produk atau entitas tertentu. Contohnya, musik, film, *game online*, selebritas, dan tokoh politik. *Fanboy K-Pop* termasuk fenomena yang normal, meski keberadaannya seolah tidak bisa lepas dari stigma. Pandangan bahwa *K-Pop* mempunyai imej musik *sweet* dan *girly* sehingga hanya cocok disukai kaum perempuan membuat *fanboy K-Pop* lebih rentan dihujat dan didiskriminasi daripada *fangirl K-Pop*. Padahal menurut Alzena Masykouri, psikolog anak dan remaja, musik bersifat universal atau tidak ada pengkhususan untuk suatu gender (R, 2019). Di samping itu, konsep maskulinitas *idol* laki-laki

Korea yang biasanya turut dianut oleh *fanboy K-Pop* berlawanan dengan standar maskulinitas hegemoni di Indonesia. Lelaki yang sangat memperhatikan penampilan, misalnya perawatan tubuh, memakai riasan, dan berbusana *trendy* dianggap kurang jantan. Pemberian stereotip aneh dan sebutan kurang menyenangkan lainnya di penolakan inilah *fanboy K-Pop* kemudian menjadi takut dan malu untuk mengekspresikan diri sesuai hobi mereka. Namun, seiring berjalannya waktu dan ketenaran *K-Pop* yang semakin menanjak, pemahaman masyarakat mengenai budaya *K-Pop* terutama *fanboy K-Pop* sedikit banyak mengalami perubahan dan terkesan “mencair”. *Fanboy K-Pop* kini berani terbuka dan tidak ragu memperlihatkan eksistensinya di kehidupan nyata maupun di media sosial.

Dari uraian di atas, penelitian ini berupaya mengkaji komentar masyarakat (netizen) Twitter terhadap *fanboy NCT* menggunakan analisis sentimen *Naïve Bayes Classifier*. Metode klasifikasi *Naïve Bayes* cukup umum di ranah penelitian *data mining* karena durasi pelaksanaan analisis data yang cepat, sistem pengoperasian mudah, dan efektifitasnya memuaskan (Prasasti et al., 2021). Perihal yang lain, kelangkaan studi terkait *fanboy K-Pop* terlebih *fanboy boygroup K-Pop* menarik peneliti untuk memilih *fanboy NCT* sebagai objek penelitian. *NCT* atau Neo Culture Technology merupakan *boygroup K-Pop* unik bentukan SM Entertainment yang menjadi salah satu artis *K-Pop* terpopuler. Penggemar *NCT* atau familier dijuluki *NCTzen* adalah sekian dari kelompok penggemar *K-Pop* terbesar di Indonesia dan dunia. Pada tahun

2020 dan 2021, akun Twitter *NCT (@NCTsmtown)* menempati posisi ke-2 di daftar artis *K-Pop* yang paling banyak dibicarakan di Twitter (Kim, 2022).

Twitter termasuk *platform media sosial* yang banyak digunakan masyarakat Indonesia. Distribusi pesan antarpenggunanya di beragam topik perbincangan begitu aktif. Peran Twitter tidak sebatas untuk berkomunikasi, tetapi juga sebagai media hiburan, sarana jual beli barang, alat pemberitaan amatir, dan menghimpun aksi sosial melalui sebuah organisasi atau perkumpulan. Twitter secara transparan menyediakan dan mengizinkan akses terprogram 5 (lima) kelompok sumber data mereka kepada perusahaan, pengembang, dan pengguna, di antaranya pengelolaan fasilitas akun dan pengguna, postingan dan balasan, *Direct Message (DM)*, iklan, dan alat penerbit.

Rumusan Masalah

Bagaimana sentimen komentar netizen Twitter terhadap *fanboy NCT*?

Tujuan Penelitian

Tujuan penelitian ini untuk mengetahui sentimen komentar netizen Twitter mengenai *fanboy NCT*.

Kajian Teori

Teori Penilaian Sosial

Teori Penilaian Sosial atau *Social Judgement Theory* dikemukakan oleh Muzafer Sherif, tokoh psikologi sosial Turki-Amerika, pada tahun 1961 yang menaruh perhatian soal bagaimana seseorang memberikan penilaian atas pernyataan yang diterimanya dan apakah

penilaian yang dibentuk berdampak pada sistem kepercayaan yang dimiliki. Teori Penilaian Sosial berlaku untuk mengamati proses seseorang mempertimbangkan pesan komunikasi. Proses penilaian pesan dinamakan persepsi sosial. Di kehidupan sosial, acuan atau referensi tersimpan berdasarkan pengalaman sebelumnya. Referensi internal atau *reference point* tersebut menentukan cara seseorang merespon suatu pesan (Morissan, 2013).

Penerapan Teori Penilaian Sosial dapat membantu penafsiran komunikasi lewat perubahan sikap melihat rentang yang ada, meliputi (Syarif, 2020):

1. Pesan yang jatuh pada rentang penerimaan (*latitude of acceptance*) akan mempermudah perubahan sikap,
2. Jika pesan terletak di rentang penolakan (*latitude of rejection*), maka perubahan sikap akan berkurang atau cenderung berlawanan, dan
3. Pesan berada dalam rentang non-komitmen atau ketidakterlibatan (*latitude of noncommitment*). Seseorang menyatakan pandangan tidak setuju, namun juga tidak menolak.

Tingkat penerimaan atau penolakan seseorang terkait suatu isu dipengaruhi sebuah variabel kunci, yaitu keterlibatan ego (*ego involvement*), yang artinya ada muatan hubungan personal dengan isu yang bersangkutan (*sense of the personal relevance of an issue*). Peran Teori Penilaian Sosial terhadap komunikasi adalah kita menyadari saat seseorang menilai suatu pesan, baik yang sifatnya menerima maupun menolak, berlandaskan referensi internal dan keterlibatan

ego. Tetapi, proses penilaian ini juga berpotensi memunculkan distorsi (penyimpangan). Seseorang mengalami distorsi ketika penilaian pesan mengandung efek pertentangan (*contrast effect*) atau efek asimilasi (*assimilation effect*). Efek kontras terjadi apabila individu menilai suatu pesan lebih jauh dan bertentangan dari sudut pandang yang semestinya, sedangkan efek asimilasi terwujud jika seseorang menilai suatu pesan menjadi lebih dekat dengan pandangannya sendiri daripada yang seharusnya (Putri, 2017).

Secara umum dalam menanggapi suatu isu, perubahan sikap lebih mudah dilakukan pada individu yang memiliki keterlibatan ego rendah. Kebalikannya, seseorang yang keterlibatan egonya tinggi akan jauh lebih sulit untuk menerima dan mengubah pandangannya. Di komunikasi, Sherif berpikir komunikator setidaknya perlu mempelajari 3 (tiga) rentang perubahan sikap dan keterlibatan ego komunikasi agar kekuatan persuasif pesan yang disampaikan bisa tepat sasaran dan sesuai harapan.

METODE PENELITIAN

Tipe Penelitian

Penelitian ini menggunakan pendekatan analisis isi kuantitatif. Analisis isi kuantitatif dimaksudkan sebagai metode penelitian yang objektif, sistematis, dan berciri kuantitatif terhadap isi pesan yang manifest atau terlihat. Analisis isi kuantitatif tidak membahas ideologi, wacana, atau makna teks yang diteliti. Pemanfaatan metode ini menyelaraskan perkembangan media, tampak dari tema penelitian yang bervariasi. Herring

(2004), di era *new media* telah dibangun perangkat lunak yang mampu mengerjakan analisis isi (Pratama et al., 2021). Analisis sentimen *Naïve Bayes* menjadi satu dari beberapa macam analisis isi yang membutuhkan *software* mesin pembelajaran dalam analisis data.

17

Jenis dan Sumber Data

Jenis penelitian ini adalah kuantitatif deskriptif. Penelitian kuantitatif deskriptif menggambarkan dan menguraikan data yang terkumpul sebagaimana adanya untuk memudahkan pemahaman (Sugiyono, 2013).

Sumber data diperoleh dari Twitter yang bersifat sumber sekunder data eksternal. Peneliti menghimpun tweet atau komentar Bahasa Indonesia berkata kunci “fanboy nct” yang diposting akun Twitter publik selama periode 3 Oktober 2022 – 3 April 2023 dengan limit data 1.500 tweet. Dari 517 tweet yang didapatkan, sampel data penelitian ini sebanyak 408 tweet (setelah melewati tahapan *cleansing* dan *preprocessing*).

Teknik Pengumpulan Data

Teknik pengumpulan data menerapkan *crawling data*, yaitu proses automasi pengumpulan data yang bersumber di internet. Penelitian ini memerlukan kode *auth_token* Twitter. Tautan *script crawling data* Twitter di Google Colaboratory (Satria, 2023).

<https://colab.research.google.com/drive/11cyL2GVKIZyY1->

[O4RNrQnsz6XtNikmNV?usp=sharing](https://colab.research.google.com/drive/11cyL2GVKIZyY1-O4RNrQnsz6XtNikmNV?usp=sharing)

Peneliti mencari buku, jurnal penelitian, skripsi, video di YouTube, artikel

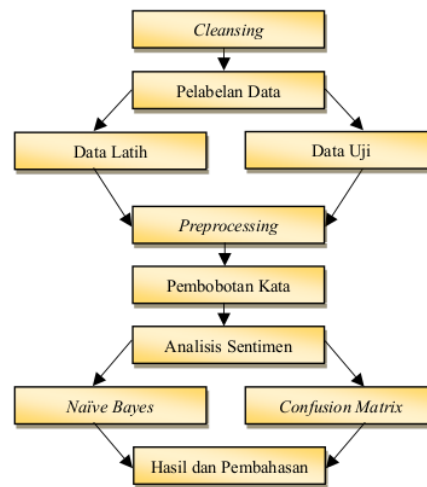
berita, situs *web*, dan dokumen pendukung lain yang relevan dengan topik *K-Pop, fanboy, NCT, Twitter, crawling data*, dan analisis sentimen *Naïve Bayes Classifier* sebagai bahan rujukan menyusun penelitian.

19

Teknik Analisis Data

Teknik analisis data penelitian ini statistik inferensial atau statistik probabilitas. Hasil analisis data statistik probabilitas diberlakukan untuk populasi. Kesimpulannya merujuk pada kemungkinan peluang (*probability*). Suatu kesimpulan mempunyai peluang kesalahan dan peluang kebenaran (kepercayaan) yang dicantumkan berupa persentase (Sugiyono, 2013). Metode analisis sentimen *Naïve Bayes* merupakan klasifikasi teks perhitungan sederhana Bayes yang menilai probabilitas suatu data berdasarkan referensi terdahulu. Peneliti menggunakan *software* RapidMiner.

Gambar 1. Alur Teknik Analisis Data



Sumber: Diolah Peneliti dari Berbagai

Sumber

Berikut penjelasan Gambar 1.

1. *Cleansing* proses membersihkan data yang tidak berguna untuk mengurangi gangguan atau *noise* saat klasifikasi sentimen, seperti menghilangkan RT, tautan, tagar, *username*, simbol-simbol, baris kosong, dan teks duplikat.
2. Pelabelan data merupakan tahap memetakan sentimen pada teks. Data yang dilabeli manual disebut *data latih* atau *data training*. Untuk data yang akan dilakukan *klasifikasi machine learning* ditulis sebagai *data uji* atau *data testing*.
3. *Preprocessing* analisis sentimen bertujuan menyeragamkan teks yang semula tidak terstruktur menjadi teks terstruktur agar data lebih layak dan siap dikerjakan mesin pembelajaran.
4. Pembobotan kata mengacu pada pemberian skor atas frekuensi atau kemunculan kata dalam dokumen teks.
5. Analisis sentimen adalah studi komputasi yang meneliti opini, sentimen, evaluasi, penilaian, sikap, dan emosi seseorang terhadap sesuatu, seperti produk, organisasi, tokoh publik, musik, peristiwa, dan sebagainya (Liu, 2012). *Naïve Bayes* termasuk model klasifikasi *supervised learning*. Evaluasi atau pengukuran kinerja klasifikasi dapat dijalankan dengan *confusion matrix* yang biasanya menggali nilai *precision*, *recall*, dan akurasi.

HASIL DAN PEMBAHASAN

Hasil analisis sentimen *Naïve Bayes Classifier* dibahas terperinci dan detail sesuai metode penelitian yang sudah dilakukan.

1. Pengumpulan Data (*Crawling Data*)

Crawling data dilaksanakan pada 18 November 2023 mengambil *tweet* atau komentar Bahasa Indonesia yang memuat kata kunci “fanboy nct” rentang waktu 3 Oktober 2022 – 3 April 2023. *File* disimpan berformat *CSV*.

Tabel 1. Hasil *Crawling Data* (Sampel)

Teks	Username
GUE SALFOK MAS MASNYA KEREN AMAT WOI FANBOY NCT SEKALINYA KELUAR MANTAPH SEKALI	haefullsun s
@nct_menfess btw salfok fanboy nya wkwwkw seru banget pariss	notbeauty s
fanboy nya nct omgggg hanbin ayo konseran bareng 🎤🎸	alvinayaa
@sweetiejis aku jg pernah ketemu sama fanboy jisung disini, awalnya cuma tau dia suka nct ternyata biasanya jisung 😊 pas ngobrol seru banget, kerasa emang beneran ngebiasin gitu 🤔	poeticdrea my

Sumber: Peneliti

Data hasil *crawling* atau dataset mentah tidak semua dimasukkan sebagai sampel data.

2. *Cleansing*

Pembersihan karakter tidak berarti di dokumen teks.

Gambar 2. *Cleansing*



Sumber: RapidMiner

Deskripsi tahapan *cleansing*.

- *Replace RT*, menghapus RT dalam teks.
- *Replace URL*, menyisihkan tautan.
- *Replace Hashtag*, menghilangkan tagar.

- *Replace Mention*, menghapus *username* dan sebutan.
- *Replace Symbol*, membuang simbol `!@#%&^&*()_+.`
- *Trim*, mengeluarkan baris teks kosong.
- *Filter Examples* bermaksud menyaring data sesuai kelompok tertentu.
- *Remove Duplicates* untuk menghilangkan teks yang sama, meski waktu *tweet* berbeda.

Tabel 2. Hasil *Cleansing*

<i>Crawling Data</i>	<i>Cleansing</i>
aku mendengar teriakan fanboy disini #LOEWExTAEYONG Taeyong 태용 [NCT] - Loewe Fall Winter 2023/2024 fashion show in Paris - 0... https://t.co/leOT3OreQx lewat @YouTube	aku mendengar teriakan fanboy disini Taeyong NCT Loewe Fall Winter 20232024 fashion show in Paris 0

Sumber: *Peneliti*

3. Pelabelan Data

Peneliti menghilangkan teks yang kalimatnya kurang bermakna, tidak sesuai konteks variabel penelitian, dan *tweet* bahasa lain pada seleksi data manual. Selanjutnya, data latih diberi label sentimen positif, negatif, dan netral sesuai pembagian skenario data.

Tabel 3. *Skenario Data Analisis Sentimen*

	Data Latih		Data Uji		Jumlah Data
	Rasio	Data	Rasio	Data	
Ske. 1	70%	286	30%	122	408
Ske. 2	80%	326	20%	82	408

Sumber: *Peneliti*

Di bawah ini sentimen komentar netizen Twitter terhadap *fanboy* NCT.

Tabel 4. Hasil Pelabelan Data

Teks	Sentimen
GUE SALFOK MAS MASNYA KEREN AMAT WOI FANBOY NCT	Positif

SEKALINYA KELUAR MANTAPH SEKALI	
lawak lo fanboy toxic minimal hapus dulu itu profil postingan nct sama komen GA pc net nya malu malu in bgt	Negatif
br kali ini liat cowo fanboy jelasin sistem nct selain doyoung	Netral

Sumber: *Peneliti*

4. Preprocessing

Preprocessing merupakan proses eliminasi dokumen teks yang mulanya tidak beraturan menjadi teks terstruktur.

Gambar 3. *Preprocessing*



Sumber: *RapidMiner*

Tahapan perlu yang dilalui dalam *preprocessing* analisis sentimen.

1. Case Folding

Menyeimbangkan karakter teks ke huruf kecil (*lower case*) untuk memastikan semua teks di dokumen memiliki format yang konsisten dan standar (Rizkina & Hasan, 2023).

Tabel 5. Hasil *Case Folding*

<i>Cleansing</i>	<i>Case Folding</i>
aku mendengar teriakan fanboy disini Taeyong NCT Loewe Fall Winter 20232024 fashion show in Paris 0	aku mendengar teriakan fanboy disini taeyong nct loewe fall winter 20232024 fashion show in paris 0

Sumber: *Peneliti*

2. Tokenization

Memisahkan setiap kata yang menyusun kalimat dengan spasi sebagai pemisahannya.

Tabel 6. Hasil *Tokenization*

<i>Case Folding</i>	<i>Tokenization</i>
aku mendengar teriakan fanboy	aku mendengar

disini Taeyong NCT Loewe Fall Winter 20232024 fashion show in paris 0	teriakan
	fanboy
	disini
	taeyong
	nct
	loewe
	fall
	winter
	20232024
	fashion
	show
	in
	paris
	0

Sumber: Peneliti

3. Stopwords Removal

Menyisihkan kata yang kurang dibutuhkan merujuk pada kamus *stopwords removal* Bahasa Indonesia yang tersedia di kaggle (Hartono, 2016).

Tabel 7. Hasil Stopwords Removal

Tokenization	Stopwords Removal
aku	-
mendengar	mendengar
teriakan	teriakan
fanboy	fanboy
disini	-
taeyong	taeyong
nct	nct
loewe	loewe
fall	fall
winter	winter
20232024	-
fashion	fashion
show	show
in	in
paris	paris
0	-

Sumber: Peneliti

4. Filter Tokens

Mengeluarkan kata yang tidak sesuai kategori. Peneliti hanya menarik kata yang tersusun dari 3 huruf sampai 25 huruf.

Tabel 8. Hasil Filter Tokens

Stopwords Removal	Filter Tokens
mendengar	mendengar

teriakan	teriakan
fanboy	fanboy
taeyong	taeyong
nct	nct
loewe	loewe
fall	fall
winter	winter
fashion	fashion
show	show
in	-
paris	paris

Sumber: Peneliti

5. Stemming

Mengubah kata yang berimbuhan menjadi kata dasar. Seperti tahapan *stopwords removal*, *stemming* memerlukan daftar kata dasar Bahasa Indonesia sebagai acuan pra-pemrosesan teks (Priawadi, 2012).

Tabel 9. Hasil Stemming

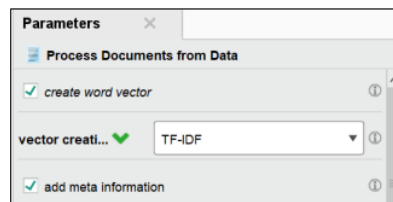
Filter Tokens	Stemming
mendengar	dengar
teriakan	teriak
fanboy	fanboy
taeyong	taeyong
nct	nct
loewe	loewe
fall	fall
winter	winter
fashion	fashion
show	show
paris	paris

Sumber: Peneliti

5. Pembobotan Kata (TF-IDF)

Pembobotan kata bertujuan mengonversi kata dalam angka.

Gambar 4. Pembobotan Kata TF-IDF



Sumber: RapidMiner

Metode TF-IDF mengukur nilai frekuensi sebuah kata pada dokumen teks. Semakin banyak suatu kata muncul di korpus data, maka bobot kata semakin tinggi.

Tabel 10. Hasil Pembobotan TF-IDF

Kata	TF-IDF
fall	0.350
fanboy	0.002
fashion	0.350
loewe	0.350
mendengar	0.350
nct	0.008
paris	0.350
show	0.350
taeyong	0.222
teriakan	0.309
winter	0.350

Sumber: Peneliti

6. Analisis Sentimen Naïve Bayes Classifier
- Naïve Bayes Classifier merupakan metode analisis sentimen yang menerapkan perhitungan probabilitas yang mengasumsikan setiap atribut bersifat bebas (independen) (Cindo et al., 2019).

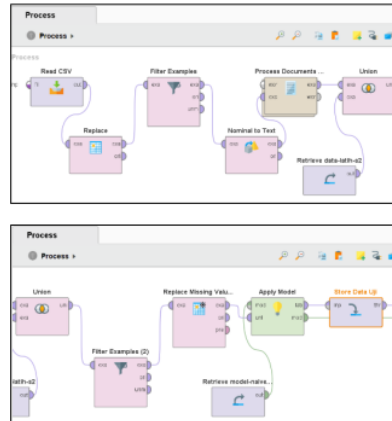
Gambar 5. Pengembangan Model



Sumber: RapidMiner

Peneliti memerlukan data latih guna mengembangkan model klasifikasi Naïve Bayes. Model klasifikasi yang selesai dibuat kemudian diimplementasikan pada data uji untuk memprediksi teks yang belum berlabel sentimen.

Gambar 6. Implementasi Model



Sumber: RapidMiner

Perhitungan Naïve Bayes ialah penerapan teorema Bayes di lingkup big data analysis. Teorema Bayes ditemukan oleh Thomas Bayes, seorang matematikawan Inggris abad ke-18 (Ghozali et al., 2023).

Tabel 11. Hasil Analisis Sentimen

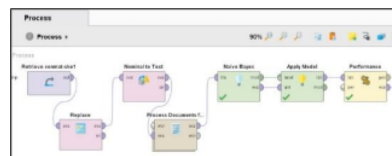
Data	Hasil Analisis Sentimen			
	Positif	Negatif	Netral	
Ske. 1	Data	287	25	96
	Persentase	70,3%	6,1%	23,5%
Ske. 2	Data	276	26	106
	Persentase	67,6%	6,4%	26 %

Sumber: Peneliti

7. Confusion Matrix

Confusion matrix atau matriks kebingungan, yaitu metode pengukuran yang digunakan untuk mengevaluasi kinerja model klasifikasi machine learning (Febriyanti, 2020).

Gambar 7. Confusion Matrix



Sumber: RapidMiner

Confusion matrix dilakukan mengikuti skenario data yang sudah diatur.

Tabel 12. Hasil *Confusion Matrix*
Skenario 1

Data		Prediksi			Recall
		Positif	Negatif	Netral	
Aktual	Positif	282	1	4	98,26%
	Negatif	0	25	0	100%
	Netral	0	2	94	97,92%
Precision		100%	89,29%	95,92%	
Mean Precision 95,07%					
Mean Recall 98,72%					
Akurasi 98,28%					

Sumber: Peneliti

Tabel 13. Hasil *Confusion Matrix*
Skenario 2

Data		Prediksi			Recall
		Positif	Negatif	Netral	
Aktual	Positif	270	3	3	97,83%
	Negatif	0	26	0	100%
	Netral	0	2	104	98,11%
Precision		100%	83,87%	97,2%	
Mean Precision 93,69%					
Mean Recall 98,65%					
Akurasi 98,04%					

Sumber: Peneliti

Keseluruhan dokumen teks yang berjumlah 408 data, skenario 1 menghasilkan sentimen positif mencapai 70,3%, negatif 6,1%, dan netral 23,5%. Pengukuran *confusion matrix* memperoleh presisi 95,07%, *recall* 98,72%, dan akurasi 98,28%. Temuan di skenario 2 kelas positif sebanyak 67,6%, negatif 6,4%, dan netral 26%. Evaluasi matriks untuk *precision* 93,69% dan *recall* 98,65%. Tingkat akurasinya 98,04%.

Bila dikaitkan dengan Teori Penilaian Sosial, penerimaan netizen Twitter terhadap *fanboy* NCT lebih besar dibandingkan penolakan dan ketidakberpihakan (netral). Pengalaman atau referensi internal dan keterlibatan ego berbeda masing-masing individu memengaruhi persepsi sosial yang menimbulkan perubahan sikap. Perubahan sikap netizen Twitter mengenai *fanboy* NCT

bisa diamati melalui pengukuran orientasi sentimen komentar mereka apakah positif, negatif, atau netral. Penelitian ini tidak mendalami pengalaman dan keterlibatan ego serta makna dibalik komentar netizen Twitter.

PENUTUP

Kesimpulan

Berdasarkan analisis data dan pembahasan di bab sebelumnya, ditarik kesimpulan bahwa hasil analisis sentiment komentar netizen Twitter terhadap *fanboy* NCT yang paling optimal terletak pada skenario 1 dengan sentimen positif 287 data (70,3%), sentimen negatif 25 data (6,1%), dan sentiment netral 96 data (23,5%). Uji evaluasi model klasifikasi nilai presisi sebesar 95,07%, *recall* 98,72%, dan akurasi 98,28%. Tingkat kesalahan klasifikasi 1,72%. Maka demikian, mayoritas netizen Twitter di Indonesia menunjukkan respons yang positif terhadap *fanboy* NCT. Komentar negatif dan netral yang jumlahnya lebih kecil membuktikan adanya variasi pandangan di kalangan netizen. Tingginya persentase performa klasifikasi mengindikasikan pemodelan *Naïve Bayes* sangat baik dan stabil dalam mengidentifikasi dan mengklasifikasikan sentiment.

Saran

Saran oleh peneliti, antara lain:

1. Saran Teoritis
Komunikasi erat berkaitan dengan perkembangan media. Maka, saran teoritis untuk penelitian selanjutnya bisa mengembangkan teori dan analisis sentimen yang mengangkat fenomena

sosial atau tema penelitian lain asalkan masih berada di cakupan disiplin ilmu komunikasi.

2. Saran Praktis

Saran praktis dari hasil analisis sentiment penelitian ini, pembaca mampu bijak dalam memahami fenomena *fanboy* NCT dan berbagai komentar netizen terhadap kelompok subkultur tersebut.

DAFTAR PUSTAKA

- Cindo, M., Rini, D. P., & Ermatita. (2019). Literatur Review: Metode Klasifikasi Pada Sentimen Analisis. *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS)*, 66–70. <https://seminar-id.com/semnas-sainteks2019.html>
- Febriyanti, A. (2020). *ANALISIS SENTIMEN PERSEPSI PENGGUNA JNE MENGGUNAKAN ALGORITMA NAÏVE BAYES CLASSIFIER* [Universitas Islam Indonesia Yogyakarta]. <http://dspace.uui.ac.id/123456789/24681>
- Ghozali, M. I., Sugiharto, W. H., & Iskandar, A. F. (2023). Analisis Sentimen Pinjaman Online Di Media Sosial Twitter Menggunakan Metode Naive Bayes. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 3(6), 1340–1348. <https://doi.org/10.30865/klik.v3i6.936>
- Haloho, L. T. M., Danadharta, I., & Kusumaningrum, H. (2023). Analisis Semiotika Representasi Soft Masculinity pada Iklan Kosmetik SOMETHINC x NCT Dream. *SEMAKOM: SEMINAR NASIONAL MAHASISWA KOMUNIKASI*, 1(1). <https://conference.untag-sby.ac.id/index.php/semakom/article/view/1906>
- Hartono, O. R. (2016). *Indonesian Stoplist*. Kaggle. <https://www.kaggle.com/datasets/oswinrh/indonesian-stoplist>
- Kim, Y. (2022). *Indonesia negara yang paling ramai membicarakan K-pop di Twitter 2021*. Blog Twitter. https://blog.twitter.com/in_id/topics/insights/2022/indonesia-negara-yang-paling-ramai-membicarakan-k-pop-di-twitter
- Liu, B. (2012). Sentiment Analysis and Opinion Mining. In *Morgan & Claypool Publishers* (Issue May). Morgan & Claypool Publishers. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.244.9480&rep=rep1&type=pdf>
- Morissan. (2013). *TEORI KOMUNIKASI Individu Hingga Masa* (Pertama). Jakarta: KENCANA PRENAMEDIA GROUP.
- Prasasti, Q. A., Ramantoko, G., & Nurhazizah, E. (2021). ANALISIS SENTIMEN MASYARAKAT TERHADAP ULASAN iOS 14 BERDASARKAN OPINI PENGGUN TWITTER MENGGUNAKAN NAÏVE BAYES CLASSIFIER. *E-Proceeding of Management*, 8(4), 3085–3094. <https://openlibrarypublications.telkomuniversity.ac.id/index.php/management/article/view/15295>
- Pratama, B. I., Illahi, A. K., Pratama, M. R., Anggraini, C., & Ari, D. P. S. (2021). *METODE ANALISIS ISI (Metode Penelitian Populer Ilmu-ilmu Sosial)* (Pertama). Malang: Unisma Press.
- Priawadi, O. (2012). *Download database kata dasar bahasa indonesia*. Priawadi. <https://www.priawadi.com/2012/11/download-database-kata-dasar-bahasa.html>
- Putri, D. K. Y. S. (2017). *TEORI KOMUNIKASI* (D. Angraeni (ed.); (Cetakan Pertama). Jakarta: Nerbitinbuku.com.
- R, N. E. (2019). *Fanboy K-Pop, Kerap Dicaci hingga Dianggap Banci*. KumparanK-POP. <https://kumparan.com/kumparank-pop/fanboy-k-pop-kerap-dicaci-hingga-dianggap-banci-1552711509998350610/full>
- Ri'aeni, I., Suci, M., Pertiwi, M., & Sugiarti, T. (2019). PENGARUH BUDAYA KOREA (K-POP) TERHADAP REMAJA DI KOTA CIREBON. *Communications*, 1(1), 1–25. <https://journal.unj.ac.id/unj/index.php/communications/article/view/9460/6798>
- Rizal, J. G. (2022). *10 Tahun Lalu, Gangnam Style Mendapatkan 1 Miliar "Views" di YouTube*. KOMPAS.Com. <https://www.kompas.com/cekfakta/read/2022/12/21/151600182/10-tahun-lalu-gangnam-style-mendapatkan-1-miliar-views-di-youtube?page=all>

- Rizkina, N. Q., & Hasan, F. N. (2023). Analisis Sentimen Komentar Netizen Terhadap Pembubaran Konser NCT 127 Menggunakan Metode Naive Bayes. *Journal of Information System Research (JOSH)*, 4(4), 1136–1144. <https://doi.org/10.47065/josh.v4i4.3803>
- Satria, H. (2023). *Cara mendapatkan data twitter yang besar (lebih dari 1500)*. Blog Helmi Satria. <https://helimisatria.com/blog/cara-mendapatkan-data-twitter-yang-besar>
- Sugiyono, P. D. (2013). *METODE PENELITIAN KUANTITATIF, KUALITATIF DAN R & D* (Kedua). Bandung: ALFABETA.
- Syarif, E. (2020). PENGARUH MEDIA SOSIAL TERHADAP SIKAP DAN PENDAPAT PEMUDA MENGENAI UJARAN KEBENCIAN. *Jurnal Common*, 3(2), 120–141. <https://doi.org/10.34010/common.v3i2.2602>
- Triadanti. (2019). *Jadi Gaya Hidup, Benarkah Fans KPop Kaya Raya atau Cuma Modal Kuota?* IDN Times. <https://www.idntimes.com/hype/entertainment/danti/jadi-gaya-hidup-benarkah-fans-kpop-kaya-raya-atau-cuma-modal-kuota?page=all>
- Won, L. J. (2023). *Hallyu fans surpass 178 million in 2022: Report*. Korea JoongAng Daily. <https://koreajoongangdaily.joins.com/2023/03/22/entertainment/kpop/Korea-Hallyu-Kpop/20230322143703130.html>

JURNAL

ORIGINALITY REPORT

13%

SIMILARITY INDEX

10%

INTERNET SOURCES

3%

PUBLICATIONS

4%

STUDENT PAPERS

PRIMARY SOURCES

1	repository.unhas.ac.id Internet Source	1%
2	ojs.unikom.ac.id Internet Source	1%
3	123dok.com Internet Source	1%
4	Rizqia Lestika Atimi, Enda Esyudha Pratama. "Implementasi Model Klasifikasi Sentimen Pada Review Produk Lazada Indonesia", Jurnal Sains dan Informatika, 2022 Publication	1%
5	id.123dok.com Internet Source	1%
6	repository.iainpare.ac.id Internet Source	1%
7	Submitted to Universitas Brawijaya Student Paper	1%
8	www.researchgate.net Internet Source	<1%

9	jurnal.untag-sby.ac.id Internet Source	<1 %
10	Submitted to Tarumanagara University Student Paper	<1 %
11	Panji Al Muqsith Prasetyo, Arief Hermawan. "Analisis sentimen twitter terhadap pemilihan presiden menggunakan algoritma naïve bayes", INFOTECH : Jurnal Informatika & Teknologi, 2023 Publication	<1 %
12	eprints.akakom.ac.id Internet Source	<1 %
13	Submitted to Universitas Trunojoyo Student Paper	<1 %
14	Submitted to Academic Library Consortium Student Paper	<1 %
15	seminar.iaii.or.id Internet Source	<1 %
16	journals.upi-yai.ac.id Internet Source	<1 %
17	repository.uksw.edu Internet Source	<1 %
18	adoc.pub Internet Source	<1 %

19	repository.stipjakarta.ac.id Internet Source	<1 %
20	Chely Aulia Misrun, Elin Haerani, Muhammad Fikry, Elvia Budianita. "Analisis sentimen komentar youtube terhadap Anies Baswedan sebagai bakal calon presiden 2024 menggunakan metode naive bayes classifier", Jurnal CoSciTech (Computer Science and Information Technology), 2023 Publication	<1 %
21	blog.twitter.com Internet Source	<1 %
22	ejurnal.seminar-id.com Internet Source	<1 %
23	covid19mena.com Internet Source	<1 %
24	docobook.com Internet Source	<1 %
25	etheses.uin-malang.ac.id Internet Source	<1 %
26	jurnal.untagsmg.ac.id Internet Source	<1 %
27	kumparan.com Internet Source	<1 %
28	www.djournals.com	

29

Rizka Awalia Kamila, Laksmi Laksmi.
"Informasi Mengenai Perempuan dalam
Berita di Kompas.com Female", Jurnal Ilmu
Informasi, Perpustakaan, dan Kearsipan, 2018

Publication

<1 %

30

Fadhilah Dwi Ananda, Yoga Pristyanto.
"Analisis Sentimen Pengguna Twitter
Terhadap Layanan Internet Provider
Menggunakan Algoritma Support Vector
Machine", MATRIK : Jurnal Manajemen, Teknik
Informatika dan Rekayasa Komputer, 2021

Publication

<1 %

Exclude quotes Off

Exclude matches Off

Exclude bibliography On