

Klasifikasi Sentimen Masyarakat Terhadap Guru PNS yang Resign Karena Lingkungan Toxic Menggunakan Metode Naive Bayes Classifier

Rendi Ibnu Syahputra¹, Elin Haerani^{*2}, Liza Afriyanti³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau

e-mail: ^{*1}12050117694 @students.uin-suska.ac.id, ²elin.haerani@uin-suska.ac.id, ³liza.afriyanti@uin-suska.ac.id

Abstrak

Kasus pengunduran diri seorang guru Pegawai Negeri Sipil (PNS) akibat lingkungan kerja yang tidak sehat atau toxic workplace menjadi sorotan publik dan ramai diperbincangkan di media sosial, memunculkan berbagai tanggapan dari masyarakat. Penelitian ini bertujuan untuk mengklasifikasikan sentimen masyarakat terhadap fenomena tersebut dengan menganalisis 1.000 komentar yang dikumpulkan secara manual dari tiga akun media sosial berita di Instagram. Komentar tersebut dilabeli secara manual oleh seorang ahli bahasa Indonesia menjadi dua kelas, yaitu 500 komentar positif dan 500 komentar negatif. Data dianalisis menggunakan metode Naive Bayes Classifier yang diawali dengan tahap text preprocessing, termasuk pembersihan teks, tokenisasi, stopword removal, dan stemming, kemudian dilanjutkan dengan pemberian bobot menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Model dikembangkan menggunakan bahasa pemrograman Python dengan library scikit-learn, dan dievaluasi menggunakan confusion matrix dengan rumus akurasi: $(\text{jumlah prediksi benar} / \text{total data}) \times 100\%$. Hasil pengujian menunjukkan bahwa model berhasil mencapai tingkat akurasi sebesar 93,81%, yang menunjukkan bahwa pendekatan klasifikasi ini efektif dalam mengidentifikasi sentimen publik. Temuan ini memberikan gambaran mendalam mengenai pandangan masyarakat terhadap isu pengunduran diri guru PNS akibat lingkungan kerja yang tidak kondusif serta menunjukkan potensi analisis sentimen sebagai alat untuk memahami opini publik dalam isu sosial.

Kata kunci: Sentimen masyarakat, guru PNS, lingkungan toxic, Naive Bayes Classifier, klasifikasi sentimen.

Abstract

The case of the resignation of a Civil Servant (PNS) teacher due to an unhealthy work environment or toxic workplace has become a public spotlight and is widely discussed on social media, giving rise to various responses from the public. This study aims to classify public sentiment towards this phenomenon by analyzing 1,000 comments collected manually from three news social media accounts on Instagram. The comments were manually labeled by an Indonesian language expert into two classes, namely 500 positive comments and 500 negative comments. The data were analyzed using the Naive Bayes Classifier method which began with the text preprocessing stage, including text cleaning, tokenization, stopword removal, and stemming, then continued with weighting using the Term Frequency-Inverse Document Frequency (TF-IDF) method. The model was developed using the Python programming language with the scikit-learn library, and evaluated using a confusion matrix with the accuracy formula: $(\text{number of correct predictions} / \text{total data}) \times 100\%$. The test results showed that the model successfully achieved an accuracy level of 93.81%, indicating that this classification approach is

effective in identifying public sentiment. These findings provide an in-depth picture of the public's views on the issue of civil servant teacher resignations due to an uncondusive work environment and demonstrate the potential of sentiment analysis as a tool for understanding public opinion on social issues.

Keywords: *Public sentiment, civil servant teachers, toxic environment, Naïve Bayes Classifier, sentiment classification.*

1. PENDAHULUAN

Guru merupakan individu memiliki peran sangat penting dalam dunia pendidikan. Selain bertugas untuk mendidik, guru juga berfungsi sebagai teladan bagi peserta didik[1]. Dalam sistem pendidikan di Indonesia, guru yang berstatus sebagai PNS umumnya mendapatkan jaminan pekerjaan yang lebih stabil, termasuk gaji tetap, tunjangan, dan pensiun. Keberadaan jaminan tersebut dianggap dapat meningkatkan kinerja guru PNS dalam menjalankan tugasnya[2].

Meskipun terdapat jaminan kesejahteraan, kinerja sebagian guru PNS masih belum memenuhi standar yang diharapkan. Lingkungan kerja yang tidak kondusif, seperti komunikasi yang buruk, konflik interpersonal, dan praktik manajemen yang tidak adil, sering kali berdampak negatif terhadap motivasi dan produktivitas guru. Sebaliknya, lingkungan kerja yang sehat dapat mendukung peningkatan kinerja, loyalitas, serta peluang pengembangan karir yang lebih baik[3].

Akhir-akhir ini, media sosial makin sering jadi tempat ngobrol soal berbagai isu, termasuk soal kerjaan guru PNS. Tidak seperti televisi atau radio yang bersifat satu arah, media sosial memungkinkan interaksi dua arah, di mana pengguna dapat memberikan komentar dan tanggapan terhadap suatu isu yang sedang berkembang [4]. Isu yang sering dibahas adalah keputusan guru PNS untuk mengundurkan diri akibat lingkungan kerja yang dianggap beracun (toxic). Instagram, sebagai salah satu platform media sosial yang populer, sering digunakan untuk menyampaikan opini terkait fenomena ini. Dengan jumlah karakter maksimal 2.500 pada komentar dan caption, Instagram menyediakan ruang diskusi yang lebih panjang dibandingkan Twitter, yang hanya memungkinkan 280 karakter. Kompleksitas teks komentar di Instagram membuatnya menjadi tantangan tersendiri dalam analisis sentimen[5].

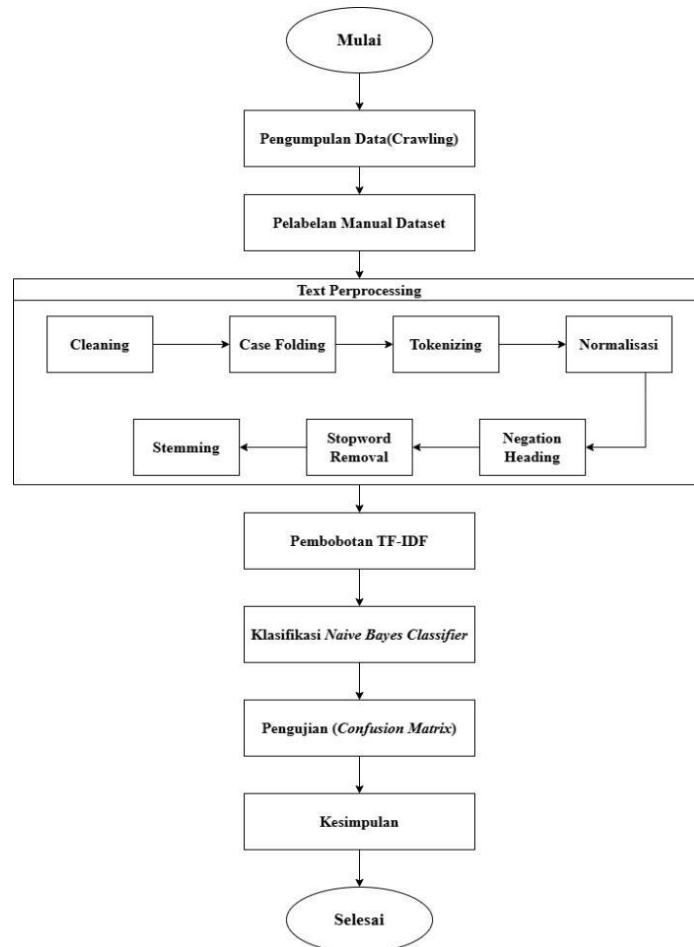
Komentar di Instagram memiliki nilai informatif yang tinggi dan dapat digunakan untuk memahami sentimen publik terhadap fenomena guru PNS yang mengundurkan diri akibat lingkungan kerja yang tidak sehat[6]. Untuk menganalisis sentimen masyarakat, data komentar dari Instagram dikumpulkan dan diproses menggunakan pendekatan *machine learning*[7]. Metode yang banyak digunakan untuk analisis sentimen adalah Multinomial Naive Bayes, yang telah terbukti efektif dalam mengklasifikasikan opini publik terhadap berbagai isu[8]. Naive Bayes Classifier dikenal mampu memberikan hasil yang akurat dalam menentukan apakah suatu opini bersifat positif atau negatif[9].

Penelitian sebelumnya menunjukkan bahwa algoritma Naive Bayes Classifier cukup efektif untuk klasifikasi sentimen dalam berbagai konteks. Salah satunya adalah studi tentang pelarangan pernikahan beda agama di Instagram, yang menggunakan 500 data positif dan 500 negatif. Studi tersebut, dengan menerapkan TfidfVectorizer—seperti dalam penelitian ini—berhasil mencapai akurasi 76,92%. Proses ini penting untuk mengubah teks tak terstruktur menjadi bentuk numerik yang dapat diproses oleh algoritma machine learning [10]. Penelitian tentang Pilpres 2024 dengan 1000 data dan skema latih-uji 90:10 menghasilkan akurasi 82% dan F1-Score 83,93% [11]. Dari 510 dataset berita yang digunakan, penelitian klasifikasi berita memperoleh akurasi tertinggi sebesar 94% pada skema pembagian 90% data latih dan 10% data uji [12].

Berdasarkan efektivitas metode yang digunakan, penelitian ini bertujuan mengklasifikasikan opini publik terkait pengunduran diri guru PNS akibat lingkungan kerja yang toxic. Dengan memanfaatkan *Naive Bayes Classifier*, penelitian ini diharapkan mampu menghasilkan klasifikasi sentimen masyarakat secara akurat terhadap fenomena tersebut.

2. METODE PENELITIAN

Penelitian memiliki enam tahap: pengumpulan data, pelabelan manual, *preprocessing text*, pembobotan kata, klasifikasi dengan Naive Bayes Classifier, dan pengujian. Alur tiap tahap ditampilkan pada Gambar 1



Gambar 1. Tahap Penelitian

2.1 Pengumpulan Data

Proses pengumpulan data mencakup peninjauan metode serta kategori yang digunakan dalam klasifikasi. Komentar yang diperoleh kemudian dikategorikan sebagai positif atau negatif. Dalam penelitian ini, sebanyak 1.000 komentar berbahasa Indonesia dikumpulkan dari postingan Instagram *Hendrabrudy*, *gurujuara.official*, dan *asninstitute* menggunakan alat *Easycomment.ai*. Data yang dihimpun masih dalam bentuk mentah dan belum diberi label.

2.2 Pelabelan Manual Dataset

Semua data yang dikumpulkan perlu diberi label secara manual. Komentar diklasifikasikan sebagai positif atau negatif berdasarkan isinya. Label positif diberikan untuk komentar yang mendukung kebijakan, seperti apresiasi, masukan konstruktif, atau respons positif

lainnya. Sementara itu, label negatif diberikan untuk komentar yang menolak, mengkritik secara tidak konstruktif, atau berpotensi memicu perpecahan, seperti provokasi dan SARA. Seorang dosen Bahasa Indonesia Roza Afifah, S.Pd, M.Hum yang memiliki 5 tahun pengalaman secara manual meninjau setiap komentar dan menetapkan label yang sesuai.

2. 3 Text Preprocessing

Preprocessing teks terdiri dari beberapa langkah untuk membersihkan data teks (menghilangkan kata) sebelum diolah [13]. Ada tujuh tahap text preprocessing yang harus dilakukan.

1. Proses membersihkan kata-kata yang tidak penting seperti simbol, tanda baca, akun, URL, hashtag, dan emoji disebut pembersihan.

2. Proses memotong teks menjadi huruf kecil yang disebut case folding.

3. Tokenizing

Proses memecah teks menjadi token (kata atau frasa kecil) untuk memudahkan analisis.

4. Normalisasi

Mengubah kata tidak baku atau salah eja menjadi bentuk yang sesuai kaidah bahasa, biasanya dengan bantuan kamus normalisasi.

5. Negation Handling

Teknik untuk mendeteksi kata bernuansa negasi dan mengubahnya menjadi lawan maknanya.

6. Stopword Removal

Menghilangkan kata-kata umum yang tidak memberikan kontribusi signifikan dalam analisis, seperti “dan”, “yang”, atau “di”.

7. Stemming

Mengkonversi kata ke bentuk dasarnya dengan menghapus imbuhan, seperti awalan atau akhiran

2. 4 Pembobotan TF-IDF

TF-IDF digunakan dalam menghitung frekuensi kata dalam dokumen dan tingkat kelangkaannya di seluruh kumpulan dokumen. TF mengukur seberapa sering kata muncul, sementara IDF menunjukkan seberapa jarang kata tersebut ditemukan [14]. Dalam penelitian ini, digunakan *TfidfVectorizer* untuk mengubah data teks menjadi vektor numerik. Proses ini memungkinkan teks tak terstruktur diproses oleh algoritma *machine learning*. Bobot setiap kata dihitung menggunakan algoritma TF-IDF yang disediakan oleh modul Scikit-learn [15].

$$TF = \frac{\text{jumlah kemunculan term dalam dokumen}}{\text{total jumlah kata dalam dokumen}} \quad (1)$$

$$IDF = \log \left(\frac{\text{Jumlah dokumen dalam korpus}}{\text{jumlah dokumen yang mengandung term}} \right) \quad (2)$$

$$w_{i,j} = tf_{i,j} \times \log \left(\frac{n}{D_{fi}} \right)^{nt} \quad (3)$$

Keterangan :

$w_{i,j}$	: nilai bobot untuk term i dalam dokumen j
$tf_{i,j}$	: jumlah kemunculan term i dalam dokumen
j	: total jumlah dokumen
D_{fi}	: banyaknya dokumen yang mengandung term i

2. 5 Klasifikasi Naive Bayes Classifier

Pengklasifikasi probabilistik yang sederhana, Pengklasifikasi Naïve Bayes menentukan probabilitas dengan memeriksa frekuensi dan kombinasi nilai dalam kumpulan data [16]. Untuk meningkatkan akurasi klasifikasi dan memperkirakan kemungkinan setiap kategori berdasarkan

karakteristik dokumen yang diuji, metode ini yang disebut Teorema Bayes dilatih menggunakan data pelatihan[17]. Rumus untuk metode klasifikasi Pengklasifikasi Naïve Bayes ditunjukkan dalam persamaan (4).

$$P(c|x) = \frac{P(c|x) P(c)}{P(x)} \quad (4)$$

Keterangan:

$P(c|x)$: Posterior probability (probabilitas kelas c setelah melihat fitur x)

$P(x|c)$: Likelihood (probabilitas mendapatkan fitur x jika diketahui kelasnya c)

$P(c)$: Prior probability (probabilitas awal kelas c)

$P(x)$: Prior probability dari fitur x secara umum

2. 6 Pengujian

Sebelum proses pengujian, data dibagi menjadi data latih dan data uji. Model klasifikasi Naive Bayes kemudian dilatih, menghasilkan model yang siap digunakan untuk memprediksi data baru [18]. Selanjutnya, model diuji pada data uji untuk menghitung nilai akurasi, ketepatan, recall, dan skor f1[19]. Tabel 1 menunjukkan matriks kekacauan yang digunakan untuk menilai kinerja klasifikasi ini.[20].

Tabel 1. Pengujian Dengan Confusion Matrix

	Actual Positive	Actual Negative
Predicted Positive	TP (True Positive)	FP (False Positive)
Predicted Negative	FN (False Negative)	TN (True Negative)

Jenis pembagian data dengan tingkat akurasi terbaik akan ditentukan oleh hasil pengujian. Rumus berikut dapat digunakan untuk menentukan skor F1, recall, akurasi, dan presisi.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

$$F1\ score = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (8)$$

3. HASIL DAN PEMBAHASAN

3. 1 Pengumpulan Data

Dengan menggunakan command exporter dari website easycomment.ai, data penelitian ini dikumpulkan dari tiga akun instagram, @hendrabrudy, @gurujuara.official, dan @asninstitute. Komentar-komentar ini berasal dari postingan yang membahas resignation seorang guru PNS karena lingkungan toxic. Baik saat mengumpulkan data maupun saat menyimpannya, data disimpan dalam format file CSV.

3. 2 Pelabelan Manual Dataset

Pelabelan data dilakukan secara manual oleh instruktur Bahasa Indonesia. Dari total 2.597 komentar yang dianalisis, dipilih 1.000 komentar untuk klasifikasi sentimen, terdiri atas 500 komentar positif dan 500 komentar negatif. Hasil pelabelan dan analisis data ditampilkan pada Gambar 2.

Sebelum dan sesudah ditinjau oleh dosen Bahasa Indonesia, Roza Afifah, S.Pd, M.Hum

Name	Text	Sentimen
titinismawa	karena itu mmg bukan jiwa seorang guru ya kan kak...guru akan lebih bijak ka	Negatif
vj.albertdaniel	@zulfaedare yes dia ngga sadar bahwa konten dia terkenal dan dikenal orang	Negatif
tlasmana12345	@rrajannah Semoga temen2nya speak up, biar nyaho 🙏	Negatif
hui_ayu	🙏 aku banget.. Rasa nya lelah.. Bener" Berasa kurang bersyukur.. 🙏 🙏	Positif
lutfisiyammiah	@martina_supit betulll..warganet banyak disetiap video pak Hendra di ambil h	Positif
hera_utomio	@nenkmutmainah mungkin kalau sesama guru tidak akan parah. Tp mungkin	Positif
akhlis.ikhlas	Kadang kadang orang lebay dan narsis berlebihan sprti ini memang tidak cocok	Negatif
1550_lela	Semangat pak guru	Positif
syerin_aa	Semangat 🙏 🙏 rencana Tuhan lebih baik dan Tuhan lebih tahu dari apa yang	Positif
puput_hindun	Suami sy guru. Nggak cocok sm orang dibilang terus. Mana ada tempat kerja y	Positif
me_ifatd	Yang begini biasanya tipekal guru yg gak sadar2 dengan kelakuan nya sendiri..	Negatif

Gambar 2. Hasil Pelabelan Manual

3. 3 Text Preprocessing

Modul Python seperti NLTK dan Pysastrawi digunakan untuk menangani dan mengatur data teks menggunakan praproses teks sebelum metode pembelajaran mesin digunakan untuk kategorisasi. Langkah- langkah dan hasil dari prosedur praproses teks ditunjukkan pada Tabel 2.

Tabel 2. Tabel Text Preprocessing

Proses	Hasil
Data Mentah	Cari topik yang bisa menginspirasi buat kita semua dan buat konten yang bisa buat kita tertawa karena keputusan sudah di depan mata harus terima dan berlapang dada
Cleaning	Cari topik yang bisa menginspirasi buat kita semua dan buat konten yang bisa buat kita tertawa karena keputusan sudah di depan mata harus terima dan berlapang dada
Case Folding	cari topik yang bisa menginspirasi buat kita semua dan buat konten yang bisa buat kita tertawa karena keputusan sudah di depan mata harus terima dan berlapang dada
Tokenizing	['cari', 'topik', 'yang', 'bisa', 'menginspirasi', 'buat', 'kita', 'semua', 'dan', 'buat', 'konten', 'yang', 'bisa', 'buat', 'kita', 'tertawa', 'karena', 'keputusan', 'sudah', 'di', 'depan', 'mata', 'harus', 'terima', 'dan', 'berlapang', 'dada']
Normalisasi	['cari', 'topik', 'yang', 'bisa', 'menginspirasi', 'buat', 'kita', 'semua', 'dan', 'buat', 'konten', 'yang', 'bisa', 'buat', 'kita', 'tertawa', 'karena', 'keputusan', 'sudah', 'di', 'depan', 'mata', 'harus', 'terima', 'dan', 'berlapang', 'dada']
Negation Handling	['cari', 'topik', 'yang', 'bisa', 'menginspirasi', 'buat', 'kita', 'semua', 'dan', 'buat', 'konten', 'yang', 'bisa', 'buat', 'kita', 'tertawa', 'karena', 'keputusan', 'sudah', 'di', 'depan', 'mata', 'harus', 'terima', 'dan', 'berlapang', 'dada']
Stopword Removal	['cari', 'topik', 'menginspirasi', 'konten', 'tertawa', 'keputusan', 'terima', 'berlapang', 'dada']
Stemming	['cari', 'topik', 'inspirasi', 'konten', 'tertawa', 'putus', 'terima', 'lapang', 'dada']

Sebelum dataset dikumpulkan dan dilabeli secara manual, seluruh tahapan *text preprocessing* diselesaikan terlebih dahulu, dengan stemming sebagai langkah terakhir. Gambar 3 menampilkan hasil text preprocessing pada seluruh dataset yang diproses menggunakan Google Colab.

	Text	Sentimen	stemming
0	Semangat terus bang, apapun keputusanmu itulah...	Positif	[semangat, bang, apa, putus, baik, buat, depan...]
1	@rozisabilillah koq pemikiran kt sama bang... ..	Negatif	[pikir, bang, laki, koten, kreator, dikontenin]
2	Bagaimana bisa seorang guru Dan laki laki mena...	Negatif	[guru, laki, laki, menang, videoin, cengeng, l...]
3	Percayalah apa yang Allah tetapkan untuk kita ...	Positif	[percaya, allah, tetap, baik, yakin, saudara]
4	yang menilai SKP kita hanya kepsek Bukan teman...	Positif	[nilai, skp, kepsek, teman, guru, allah, ta, a...]
...	...	...	...
995	Semangat ya..	Positif	[semangat]
996	Semangat kak 🙏🔥Pernah diposisi ini, 2 kali m...	Positif	[semangat, kak, posisi, kunci, nikmat, hidup, ...]
997	jangan melihat dari sisi lain pak, lihatlah ta...	Positif	[sisi, lihat, tangan, butuh, ulur, tangan, bim...]
998	sabar yah kak	Positif	[sabar, kak]
999	Mental bapak ini memang rapuh sepertinya	Negatif	[mental, rapuh]

Gambar 3. Hasil Text Preprocessing Seluruh Data Set

Proses *preprocessing text* menghasilkan data dikelompokkan berdasarkan frekuensi kata. Kata-kata yang paling sering muncul divisualisasikan dalam wordcloud sentimen positif dan negatif untuk menggambarkan topik serta nuansa utama dalam data.



Gambar 4. WordCloud Komentar Positif dan negatif

3. 4 Pembobotan TF-IDF

Proses pembobotan menggunakan pustaka Python TfidfVectorizer dari Scikit-learn. Gambar 5 menunjukkan bobot TF-IDF setiap kata, dengan bobot tertinggi menandakan kata yang paling penting dalam data.

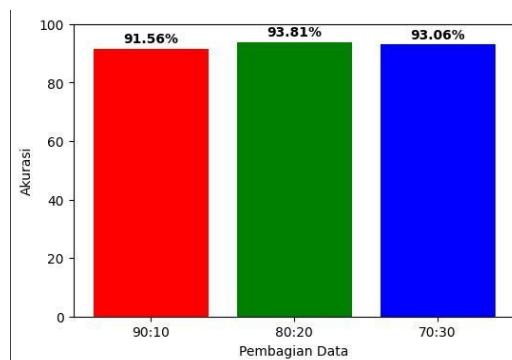
	aamiinn	abadi	abai	abang	abaykan	abdi	abdulsomad	abi	about	acting	...	year	yekaan	yes	your	yra	yuk	zaman	zholimin	zikir	zona
0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
995	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
996	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
997	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
998	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
999	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

1000 rows x 1876 columns

Gambar 5. Data Frame Pembobotan TF-IDF

3. 5 Klasifikasi Naive Bayes Classifier

Tiga pengujian Naïve Bayes dilakukan dengan pembagian data berbeda, dan akurasi tertinggi 93,81% diperoleh pada skema 80% pelatihan dan 20% pengujian.



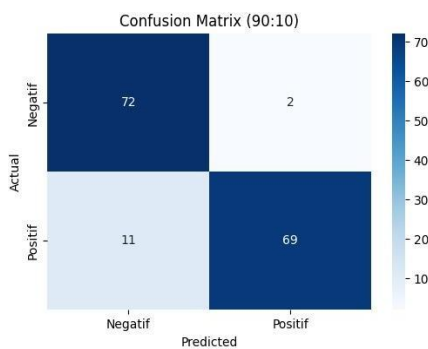
Gambar 6. Klasifikasi Pada Setiap Percobaan

3. 6 Pengujian

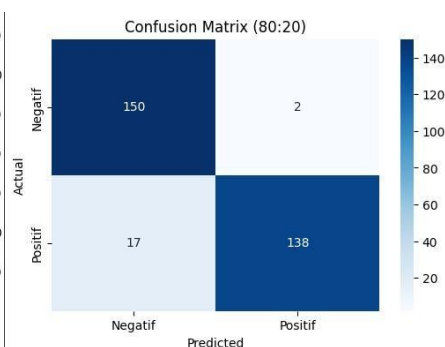
Untuk mengevaluasi efektivitas algoritma, 1.000 data dibagi menjadi set pelatihan dan pengujian dengan tiga skema berbeda: 90%–10%, 80%–20%, dan 70%–30%. Kinerja algoritma dinilai menggunakan matriks kebingungan yang membandingkan prediksi dengan label asli. Hasil evaluasi ditampilkan pada Gambar 7, 8, dan 9.

Tabel 3. Menunjukkan true dan false pada confusion matrix

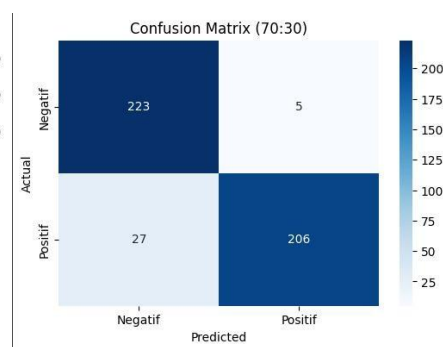
True Negative	False Negative
False Negative	True Positive



Gambar 9. Confusion Matrix Pada Pembagian Data 90%:10%



Gambar 8. Confusion Matrix Pada Pembagian Data 80%:20%



Gambar 7. Confusion Matrix Pada Pembagian Data 70%:30%

Tabel 3 menyajikan hasil pengujian untuk tiap skema pembagian data, berdasarkan matriks kebingungan pada Gambar 7, 8, dan 9, dengan perhitungan menggunakan rumus yang telah dijelaskan sebelumnya.

Tabel 4. Hasil Confusion Matrix

<i>Split Data</i>	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
90%:10%	91.56%	97.18%	86.25%	91.39%
80%:20%	93.81%	98.57%	89.03%	93.56%
70%:30%	93.06%	97.63%	88.41%	92.79%

Algoritma ini bekerja paling baik ketika skenario pemisahan data uji adalah 20%. Akurasi, presisi, recall, dan f1_score masing-masing adalah 93,81%, 98,57%, 89,03%, dan 93,56%.

4. KESIMPULAN

Penelitian ini menganalisis 1.000 komentar Instagram terkait pengunduran diri guru PNS akibat lingkungan kerja yang toxic, terdiri dari 500 komentar positif dan 500 komentar negatif. Data dikumpulkan dari tiga akun media sosial berita dan diberi label secara manual oleh seorang ahli bahasa Indonesia. Pengujian dilakukan menggunakan metode *Naïve Bayes Classifier* dengan skema pembagian data 80% untuk pelatihan dan 20% untuk pengujian. Hasil pengujian menunjukkan bahwa model ini mencapai akurasi sebesar 93,81%, precision 98,57%, recall 89,03%, dan F1-score 93,56%. Berdasarkan hasil tersebut, dapat disimpulkan bahwa metode *Naïve Bayes Classifier* memiliki tingkat keakuratan yang tinggi dalam mengklasifikasikan sentimen masyarakat terhadap fenomena pengunduran diri guru PNS karena lingkungan kerja yang toxic.

5. SARAN

Pada penelitian ini didapatkan komentar sebanyak 1000 dengan 500 positif dan 500 negatif pada komentar di sosial media instagram, kedepan nya pengujian ini diharapkan bisa menggunakan metode yang lain juga. Harapannya ada mahasiswa/i lain yang bisa menyempurnakan hasil dari yang saya buat, dan semoga kedepannya bisa lebih banyak metode yang digunakan dalam permasalahan ini.

DAFTAR PUSTAKA

- [1] Yayan Ansori Pernanda, "Perbedaan Kinerja Guru Honorer Dengan Pegawai Negeri Sipil," *J. Sci. MANDALIKA e-ISSN 2745-5955 | p-ISSN 2809-0543*, vol. 2, no. 1, pp. 37–44, 2021, doi: 10.36312/10.36312/vol2iss1pp37-44.
- [2] A. Tahun, S. Meifalanfa, F. Pradana, and R. P. Ernanda, "Perbedaan Kinerja Guru PNS dan Non-PNS : Fakta atau Stereotip ?," vol. 2, no. 1, pp. 1–8, 2025.
- [3] R. I. Maulani, N. A. Sahira, and M. Ilham, "Analisis Statistik Deskriptif Pada Lingkungan Kerja Toxic Pada Produktivitas Karyawan," vol. 2, pp. 24–38, 2025.
- [4] K. A. Nugraha, "Analisis Sentimen Berbasis Emoticon pada Komentar Instagram Bahasa Indonesia Menggunakan Naïve Bayes," *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 3, pp. 715–721, 2021, doi: 10.28932/jutisi.v7i3.4094.
- [5] A. Muhaddisi, B. N. Prastowo, and D. U. Kusumaning Putri, "Sentiment Analysis With Sarcasm Detection On Politician's Instagram," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 15, no. 4, p. 349, 2021, doi: 10.22146/ijccs.66375.
- [6] R. L. Musyarofah, E. U. Utami, and S. R. Raharjo, "Analisis Komentar Potensial pada Social Commerce Instagram Menggunakan TF-IDF," *J. Eksplora Inform.*, vol. 9, no. 2, pp. 130–139, 2020, doi: 10.30864/eksplora.v9i2.360.
- [7] F. Apri Wenando, R. Hayami, S. Soni, A. Fitria, and D. Shifana, "Sentimen Analisis Masyarakat terhadap Kasus Penembakan Brigadir J Menggunakan Algoritma Naïve Bayes Classifier," *J. CoSciTech (Computer Sci. Inf. Technol.*, vol. 4, no. 2, pp. 484–490, 2023, doi: 10.37859/coscitech.v4i2.5686.
- [8] K. Kevin, M. Enjeli, and A. Wijaya, "Analisis Sentimen Penggunaan Aplikasi Kinemaster Menggunakan Metode Naive Bayes," *J. Ilm. Comput. Sci.*, vol. 2, no. 2, pp. 89–98, 2024, doi: 10.58602/jics.v2i2.24.
- [9] N. P. G. Naraswati, R. Nooraeni, D. C. Rosmilda, D. Desinta, F. Khairi, and R. Damaiyanti, "Analisis Sentimen Publik dari Twitter Tentang Kebijakan Penanganan Covid-19 di Indonesia dengan Naive Bayes Classification," *Sistemasi*, vol. 10, no. 1, p.

- 222, 2021, doi: 10.32520/stmsi.v10i1.1179.
- [10] Muhammad Rizki Syafapri, Elin Haerani, Iwan Iskandar, and Liza Afriyanti, "Klasifikasi sentimen terhadap larangan pernikahan beda agama menggunakan metode Naive Bayes Classifier," *J. CoSciTech (Computer Sci. Inf. Technol.*, vol. 5, no. 1, pp. 10–18, 2024, doi: 10.37859/coscitech.v5i1.6889.
- [11] rahayu deny danar dan alvi furwanti Alwie, A. B. Prasetio, R. Andespa, P. N. Lhokseumawe, and K. Pengantar, "Tugas Akhir Tugas Akhir," *J. Ekon. Vol. 18, Nomor 1 Maret201*, vol. 2, no. 1, pp. 41–49, 2020.
- [12] S. M. Habib, E. Haerani, S. K. Gusti, and S. Ramadhani, "Klasifikasi Berita Menggunakan Metode Naïve Bayes Classifier," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 2, pp. 248–258, 2022, doi: 10.32672/jnkti.v5i2.4191.
- [13] R. Azhar, A. Surahman, and C. Juliane, "Analisis Sentimen Terhadap Cryptocurrency Berbasis Python TextBlob Menggunakan Algoritma Naïve Bayes," *J. Sains Komput. Inform. (J-SAKTI*, vol. 6, no. 1, pp. 267–281, 2022.
- [14] Y. Nurtikasari, Syariful Alam, and Teguh Iman Hermanto, "Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes," *INSOLOGI J. Sains dan Teknol.*, vol. 1, no. 4, pp. 411–423, 2022, doi: 10.55123/insologi.v1i4.770.
- [15] K. Kritik, T. Kepala, P. Pada, K. Di, M. Metode, and N. Bayes, "Menggunakan Metode Naive Bayes Classifier Tugas ' Akhir," 2021.
- [16] H. P. Doloksaribu and Yusran Timur Samuel, "Komparasi Algoritma Data Mining Untuk Analisis Sentimen Aplikasi Pedulilindungi," *J. Teknol. Inf. J. Keilmuan dan Apl. Bid. Tek. Inform.*, vol. 16, no. 1, pp. 1–11, 2022, doi: 10.47111/jti.v16i1.3747.
- [17] P. Arsi, B. A. Kusuma, and A. Nurhakim, "Analisis Sentimen Pindah Ibu Kota Berbasis Naive Bayes Classifier," *J. Inform. Upgris*, vol. 7, no. 1, pp. 1–6, 2021, doi: 10.26877/jiu.v7i1.7636.
- [18] D. Normawati and S. A. Prayogi, "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter," *J. Sains Komput. Inform. (J-SAKTI*, vol. 5, no. 2, pp. 697–711, 2021.
- [19] A. Nurian, "Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naïve Bayes," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3s1, pp. 829–835, 2023, doi: 10.23960/jitet.v11i3s1.3348.
- [20] A. Anjani, A. Chamid, and A. Murti, "Analisis Sentimen Kaum LGBT pada Media Sosial Twitter Menggunakan Algoritma Naïve Bayes," *JTINFO J. Tek. Inform.*, vol. 1, no. 2, pp. 1–8, 2022, [Online]. Available: <https://journal.unisnu.ac.id/JTINFO/article/view/259>