

Analisis Sentimen terhadap Kenaikan Gaji Guru Honorer Menggunakan Algoritma Naïve Bayes

Thalia Meidiyanti Fadli^{*1}, Ajeng Savitri Puspaningrum²

¹Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia, Indonesia

²Teknik Komputer, Fakultas Teknik dan Ilmu Komputer, Universitas Teknokrat Indonesia, Indonesia

Email: thalia_meidiyanti_fadli@teknokrat.ac.id, ajeng.savitri@teknokrat.ac.id

Abstrak

Kesejahteraan guru honorer menjadi isu penting dalam dunia pendidikan, salah satunya terkait kenaikan gaji. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap kenaikan gaji guru honorer menggunakan model Naïve Bayes pada dataset Twitter. Data yang digunakan diperoleh melalui crawling dari platform media sosial Twitter dengan kata kunci "kenaikan gaji guru". Proses preprocessing dilakukan meliputi cleaning, case folding, normalisasi, tokenisasi, stopword removal, dan stemming. Data dikategorikan menjadi dua kelas, yaitu sentimen positif dan negatif. Hasil analisis menunjukkan bahwa dari 3091 tweet, 88,35% termasuk dalam kategori sentimen positif, sementara 11,65% menunjukkan sentimen negatif. Model Bernoulli Naïve Bayes terbukti lebih efektif dalam menangani data tidak seimbang, dengan akurasi sebesar 87,86%. Meskipun Multinomial Naïve Bayes memiliki akurasi yang lebih tinggi yaitu 89,65%, model ini cenderung bias terhadap sentimen positif. Sebaliknya, model Gaussian Naïve Bayes menunjukkan akurasi terendah sebesar 65,53%. Penelitian ini memberikan wawasan yang penting bagi pengambil kebijakan untuk mengevaluasi opini publik terkait kesejahteraan guru honorer.

Kata Kunci: *analisis sentimen, guru honorer, kenaikan gaji, media sosial, naïve bayes, twitter.*

SENTIMENT ANALYSIS TOWARDS THE INCREASE IN HONORARY TEACHERS' SALARY USING THE NAÏVE BAYES ALGORITHM

Abstract

The welfare of honorary teachers is an important issue in the world of education, one of which is related to salary increases. This study aims to analyse public sentiment towards the salary increase for honorary teachers using the Naïve Bayes model on Twitter datasets. The data used was obtained through crawling from the Twitter social media platform with the keyword 'teacher salary increase'. The preprocessing process includes cleaning, case folding, normalisation, tokenisation, stopword removal, and stemming. The data was categorised into two classes, namely positive and negative sentiments. The analysis results show that out of 3091 tweets, 88.35% belong to the positive sentiment category, while 11.65% show negative sentiment. The Bernoulli Naïve Bayes model proved to be more effective in handling unbalanced data, with an accuracy of 87.86%. Although Multinomial Naïve Bayes has a higher accuracy of 89.65%, it tends to be biased towards positive sentiments. In contrast, the Gaussian Naïve Bayes model showed the lowest accuracy of 65.53%. This research provides important insights for policy makers to evaluate public opinion related to the welfare of honorary teachers.

Keywords: *sentiment analysis, honorary teachers, salary increase, social media, naïve bayes, twitter.*

1. PENDAHULUAN

Guru adalah salah satu elemen yang paling vital dalam bidang pendidikan. Pekerjaan guru dibagi ke dalam beberapa kategori, salah satunya adalah guru honorer. Guru yang berstatus honorer menerima bayaran secara sukarela dan bahkan sering kali di bawah tingkat upah minimum [1]. Menurut Kementerian Pendidikan dan Kebudayaan, jumlah guru honorer di Indonesia mencapai lebih dari 700.000 orang, dengan gaji bulanan rata-rata berkisar antara Rp500.000 juta hingga Rp1 juta.

Studi yang dilakukan oleh Bennell dan Akyeampong (2007) menunjukkan bahwa rendahnya tingkat gaji guru di negara berkembang merupakan salah satu faktor utama yang menyebabkan rendahnya motivasi dan produktivitas mereka. Dengan adanya peningkatan gaji, para guru dapat lebih berkonsentrasi pada tugas mereka

tanpa harus mencari pekerjaan sampingan. Situasi ini sangat relevan di Indonesia, di mana banyak guru honorer terpaksa mengambil pekerjaan paruh waktu di sektor lain untuk memenuhi kebutuhan hidup mereka [2].

Kebutuhan nasional untuk sumber daya manusia berkualitas akan dipenuhi oleh guru yang berkualitas. Namun, harapan tersebut tidak akan tercapai jika berbagai masalah pendidikan terus ditemukan. Ketidakadilan dalam dunia kerja bagi guru honorer adalah salah satu masalah di dunia pendidikan. Kehidupan mereka jauh dari kata sejahtera [3]. Kesejahteraan guru honorer telah menjadi isu yang sering memicu perdebatan, baik di tingkat pemerintahan maupun masyarakat umum [4]. Salah satu platform di mana masyarakat sering menyuarakan pendapatnya adalah media sosial, khususnya Twitter, yang sekarang dikenal sebagai X. Tweet terkait isu kesejahteraan guru, terutama tentang kenaikan gaji guru honorer telah menjadi sorotan, seperti yang dicuit oleh akun @MataNajwa di akun twitter nya dengan memberikan caption di tweet nya: “Presiden Prabowo Subianto menangis saat berpidato di hadapan para guru di Puncak Hari Guru Nasional 2024 yang diadakan di Jakarta Timur (28/11/2024).”

Twitter berfungsi sebagai platform media sosial terkemuka yang umum digunakan untuk mengukur opini publik. Analisis data dari Twitter mengungkap beberapa istilah yang sering muncul dalam tweet terkait isu kenaikan gaji guru honorer. Kata-kata yang paling umum digunakan dalam tweet meliputi "naikkan gaji guru", "kenaikan gaji guru honorer" dan banyak istilah lain yang juga muncul dalam tweet [5]. Salah satu cara untuk memahami opini masyarakat tentang topik ini adalah dengan melakukan analisis sentimen. Analisis sentimen merupakan langkah untuk mengumpulkan, mengidentifikasi, serta menilai pendapat, perspektif, atau perasaan yang terdapat dalam teks, seperti ulasan, komentar atau unggahan di media sosial. Tujuan utama dari analisis sentiment ini diangkat adalah untuk mengetahui bagaimana orang merasa atau berpendapat mengenai isu kenaikan gaji guru honorer [6].

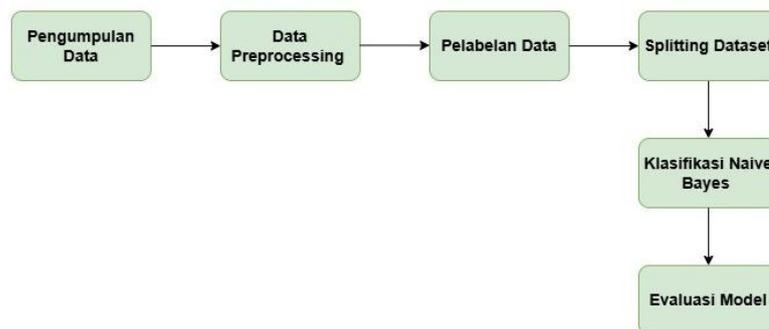
Penelitian sebelumnya yang dilakukan oleh Andriyani dkk. (2023) mengungkapkan bahwa analisis sentimen di Twitter dapat memberikan wawasan mengenai persepsi publik terhadap isu penghapusan tenaga honorer . Penelitian ini menggunakan metode Naïve Bayes Classifier untuk mengklasifikasikan sentimen dalam tweet terkait topik tersebut. Hasil penelitian menunjukkan bahwa analisis sentimen di media sosial seperti Twitter dapat mencerminkan opini publik yang berkaitan dengan kesejahteraan guru honorer, termasuk isu kenaikan gaji mereka [7].

Dalam penelitian ini, dilakukan pemodelan topik yang terdapat pada kelompok sentimen pengguna Twitter mengenai kenaikan gaji guru honorer dengan memanfaatkan Naïve Bayes untuk mengelompokkan komentar dari masyarakat menjadi positif atau negatif [8]. Alasan utama Naive Bayes dipilih sebagai metode yang efektif untuk analisis sentimen menggunakan metode NBC karena tersedia data set yang jumlahnya besar, sehingga diperlukan metode yang dapat melakukan klasifikasi dengan cepat dan memberikan tingkat akurasi yang tinggi dalam analisis [9]. Berdasarkan penelitian yang dilakukan oleh A. S. Rahayu dan A. Fauzi (2019) berjudul Komparasi Algoritma Naïve Bayes dan Support Vector Machine (SVM) dalam Analisis Sentimen Spotify, ditemukan bahwa dalam analisis sentimen terhadap 1500 ulasan Spotify, algoritma Naïve Bayes menunjukkan performa terbaik dengan tingkat akurasi sebesar 86,4%. Di sisi lain, algoritma Support Vector Machine (SVM) hanya mencapai akurasi sebesar 84% [10]. Berdasarkan alasan ini, Naive Bayes dipilih karena keseimbangan yang baik antara kecepatan, akurasi, dan kemudahan implementasi dalam analisis sentimen terhadap kenaikan gaji guru honorer. Adapun manfaat dari penelitian ini adalah memberikan gambaran umum tentang persepsi masyarakat terkait isu kenaikan gaji guru honorer yang dapat digunakan sebagai bahan pertimbangan dalam pengambilan kebijakan yang berkaitan dengan kesejahteraan guru honorer di Indonesia.

Penelitian ini bertujuan untuk mengevaluasi sentimen masyarakat terkait peningkatan gaji guru honorer dengan memanfaatkan algoritma Naïve Bayes. Melalui analisis data yang diperoleh dari Twitter, diharapkan penelitian ini dapat memberikan informasi yang berguna bagi para pembuat kebijakan dalam mempertimbangkan kesejahteraan guru honorer saat merumuskan kebijakan di masa yang akan datang.

2. METODE PENELITIAN

Ada 5 tahapan dalam proses penelitian ini, gambar 1 merupakan langkah pertama dalam mengumpulkan data yang akan dianalisis. Setelah data terkumpul, preprocessing data dilakukan untuk mempersiapkan data sebelum dianalisis. Selanjutnya, pada tahap pelabelan data, setiap teks yang telah melalui tahap pembersihan akan diberi label sesuai dengan kategori tertentu, seperti positif dan negatif. Data yang telah diberi label kemudian dibagi menjadi dua bagian, yaitu data uji dan data latih. Pada tahap NBC diterapkan untuk mengklasifikasikan data berdasarkan probabilitas.



Gambar 1. Tahap Proses Penelitian

2.1. Pengumpulan Data

Penelitian ini dimulai dengan tahap pengumpulan data yang diperoleh melalui proses crawling di platform media sosial Twitter menggunakan bahasa pemrograman Python [11]. Python merupakan bahasa pemrograman yang bersifat interpretatif, sederhana untuk dipahami, dan dapat digunakan di berbagai sistem dengan penekanan khusus pada kejelasan kode [12]. Kata kunci yang diterapkan dalam penelitian ini meliputi "kenaikan gaji guru", "guru honorer", dan "naikkan gaji guru". Data yang diperoleh mencakup tweet yang dipublikasikan dalam rentang waktu dari 1 Januari 2024 hingga 31 Desember 2024, dengan tujuan untuk menangkap pandangan masyarakat secara lebih komprehensif dan sesuai dengan kondisi terkini. Jumlah total tweet yang dikumpulkan sebanyak 4.172 tweet. Setiap tweet yang diambil diproses untuk dibersihkan dari elemen-elemen yang tidak relevan seperti URL, emotikon, dan simbol-simbol lainnya. Data hasil crawling kemudian disimpan dalam format CSV untuk memudahkan proses analisis selanjutnya. Proses crawling ini memungkinkan untuk mengumpulkan berbagai macam tweet secara cepat dan efisien, sehingga data yang diperoleh cukup besar dan representatif untuk dilakukan analisis sentimen.

2.2. Preprocessing Data

Setelah memperoleh data menggunakan metode crawling, langkah berikutnya yang dilakukan adalah preprocessing data. Preprocessing bertujuan untuk membersihkan data dari *noise*. Dalam tahap ini terdapat beberapa langkah yang akan diterapkan pada data yang ada, yaitu proses cleaning, proses case folding, normalization, tokenization, stopword removal dan stemming data [13].

1. Cleaning

Pembersihan data merupakan proses yang melibatkan penghapusan data yang dianggap tidak relevan, tidak lengkap, tidak akurat, atau memiliki format yang salah dalam basis data. Tujuan dari kegiatan ini adalah untuk menghasilkan data yang berkualitas tinggi [14]. Tabel 1 hasil contoh dari cleaning:

Tabel 1. *Cleaning*

Tweet	Cleaning
"Kenaikan gai guru sangat diharapkan! https://t.co/xyz123 "	"Kenaikan gaji guru sangat diharapkan!"

2. Case Folding

Proses case folding adalah langkah yang dilakukan untuk mengubah huruf kapital menjadi huruf kecil. Tujuan dari tahap ini adalah untuk mempermudah proses berikutnya [15]. Tabel 2 hasil contoh dari case folding:

Tabel 2. *Case Folding*

Tweet	Case Folding
"Guru honorer diangkat jadi ASN PPPK"	"guru honorer diangkat jadi asn pppk"

3. Normalization

Teknik normalisasi adalah proses yang mengubah data ke dalam format yang lebih baku. Normalisasi sering diterapkan untuk merubah teks menjadi bentuk yang lebih konsisten sebelum dilakukan analisis. Proses ini dapat meningkatkan akurasi dalam analisis teks serta mempermudah pengolahan data [16]. Tabel 3 hasil contoh dari *normalization*:

Tabel 3. *Normalization*

Tweet	Normalization
"Guru honorer nggk naik gaji"	"Guru honorer tidak naik gaji"

4. Tokenization

Tokenisasi adalah proses yang bertujuan untuk memecah teks menjadi kata-kata [17]. Tabel 4 hasil contoh dari tokenization:

Tabel 4. *Tokenization*

Tweet	Case Folding
Pemerintah telah mengumumkan rencana untuk meningkatkan gaji para guru honorer yang telah mengabdikan selama bertahun-tahun dan memberikan kontribusi signifikan bagi dunia pendidikan di Indonesia langkah ini diambil sebagai bentuk penghargaan atas dedikasi dan kerja keras yang telah mereka tunjukkan	'Pemerintah', 'telah', 'mengumumkan', 'rencana', 'untuk', 'meningkatkan', 'gaji', 'para', 'guru', 'honor', 'yang', 'telah', 'mengabdikan', 'selama', 'bertahun-tahun', 'dan', 'memberikan', 'kontribusi', 'signifikan', 'bagi', 'dunia', 'pendidikan', 'di', 'Indonesia', 'langkah', 'ini', 'diambil', 'sebagai', 'bentuk', 'penghargaan', 'atas', 'dedikasi', 'dan', 'kerja', 'keras', 'yang', 'telah', 'mereka', 'tunjukkan'

5. Stopword Removal

Stopword Removal adalah langkah untuk menghilangkan kata-kata yang dianggap tidak bermakna (misalnya: ini, itu, mau, yang). Proses ini memanfaatkan kamus stopwords dalam Bahasa Indonesia [18]. Tabel 5 hasil contoh dari stopwords removal:

Tabel 5. *Stopword Removal*

Tweet	Stopword Removal
"Para guru honorer di sekolah tersebut tengah berusaha untuk memperoleh peningkatan gaji yang sesuai"	"Guru honorer sekolah tengah berusaha memperoleh peningkatan gaji sesuai"

6. Stemming Data

Stemming merupakan salah satu metode untuk memperoleh kata dasar dari suatu kata dalam kalimat dengan cara memisahkan kata tersebut dari imbuhan yang terdapat di awal maupun di akhir [19]. Tabel 6 hasil contoh dari stemming data:

Tabel 6. *Stemming Data*

Tweet	Stemming Data
"Sudah naik gaji pokoknya"	"sudah naik gaji pokok"

2.3. Pelabelan Data

Pada tahap ini, peneliti menerapkan proses pelabelan data menggunakan model otomatis untuk mengklasifikasikan tweet yang mengandung sentimen positif atau negatif [20]. Proses ini dilakukan dengan memanfaatkan model pembelajaran mesin (machine learning) yang dilatih untuk mengenali pola sentimen berdasarkan dataset yang sudah diberi label secara manual. Pelabelan data bertujuan untuk memberikan makna atau konteks pada data yang belum diproses, sehingga algoritma bisa memahami pola dan keterkaitan antara data dan label yang diberikan dan memberikan pelabelan positif dan negatif. Tabel 7 hasil contoh dari pelabelan data.

Tabel 7. Pelabelan Data

Tweet	Sentimen
kecewa narasi gaji guru hadap serius antri ppg amanat undang selesai beliau tugas ppg daljab selesai bayar hutang.	Negatif
presiden prabowo subianto umum naik gaji guru umum sejahtera guru tingkat prabowo acara puncak ingat guru nasional.	Positif

2.4. Splitting Dataset

Setelah proses pelabelan selesai, data akan dibagi menjadi dua bagian, yaitu 80% untuk data latih dan 20% untuk data uji, dengan perbandingan 80:20 [19].

2.5. Klasifikasi Naive Bayes

Naive Bayes adalah suatu metode klasifikasi yang memanfaatkan probabilitas sederhana, yang didasarkan pada Teorema Bayes dan memiliki asumsi ketidaktergantungan yang signifikan antara setiap kondisi atau peristiwa [21].

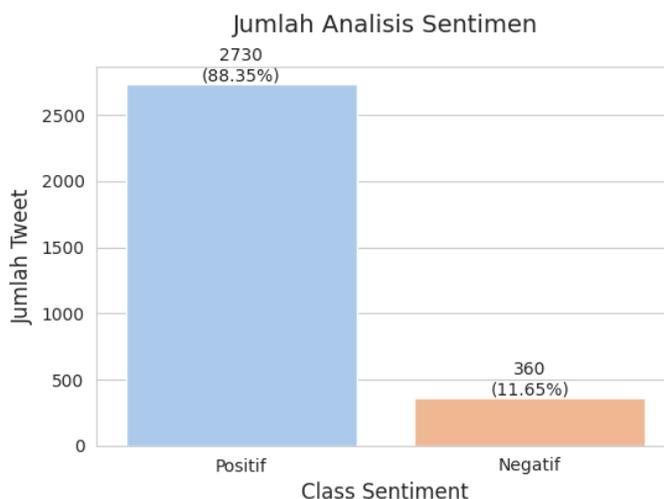
$$P(C \mid X) = (P(X \mid C) \times P(C)) \quad (1)$$

Deskripsi:

- X : Kelas data yang tidak diketahui
- C : Data Hipotesis
- $P(C \mid X)$: Probabilitas hipotesis H kondisi X
- $P(C)$: Probabilitas Hipotesis H
- $P(X \mid C)$: Probabilitas X berdasarkan kondisi dalam hipotesis H
- $P(X)$: Probabilitas X

3. HASIL DAN PEMBAHASAN

3.1. Dataset



Gambar 2. Kelas Perbandingan Analisis Sentimen Positif dan Negatif

Setelah melakukan preprocessing dan pelabelan data pada 3091 tweet yang telah dikumpulkan, hasil analisis sentimen menunjukkan bahwa 2730 tweet (88,35%) tergolong dalam kategori sentimen positif, sedangkan 360 tweet (11,65%) termasuk dalam kategori sentimen negatif. Berikut merupakan visualisasi dari perbandingan analisis sentiment positif dan negatif.

3.2. Tahap Visualisasi Word Cloud

Word cloud merupakan representasi visual yang terdiri dari kata-kata yang diatur sedemikian rupa, di mana ukuran setiap kata menunjukkan frekuensi kemunculannya dalam teks. Kata yang lebih sering muncul akan memiliki ukuran yang lebih besar dalam word cloud tersebut.. Tujuan dari penggunaan Word Cloud adalah untuk menyajikan data secara visual, sehingga memudahkan dalam mengidentifikasi istilah yang paling sering muncul.

Gambar 3 memperlihatkan hasil word cloud yang dihasilkan dari data opini dengan sentimen positif. Word cloud ini didominasi oleh kata-kata seperti gaji guru, naik gaji, guru honorer dan guru asn.

Gambar 4 memperlihatkan hasil word cloud yang dihasilkan dari data opini dengan sentimen negatif. Word cloud ini didominasi oleh kata-kata seperti gaji guru, naik gaji, guru honorer, prabowo, dan gaji jabat.



Gambar 1 Word Cloud Sentimen Positif



Gambar 2. Word Cloud Sentimen Negatif

3.3. Tahap Evaluasi Model

Tahap evaluasi model dilakukan dengan memanfaatkan data uji yang terpisah dari data pelatihan. Data uji berfungsi untuk menilai kinerja model dalam mengklasifikasikan sentimen tweet yang belum pernah dianalisis sebelumnya. Akurasi adalah metrik yang menggambarkan proporsi prediksi yang benar dibandingkan dengan total prediksi. Presisi mengukur proporsi prediksi positif yang benar dari keseluruhan prediksi positif. Recall mengukur proporsi prediksi positif yang benar dari total data positif yang sebenarnya. F1-score adalah rata-rata harmonis antara presisi dan recall.

Tabel 8. Hasil Nilai Akurasi

Model	Akurasi
GaussianNB	65,53%
MultinomialNB	89,65%
BernoulliNB	87,86%

Pada tabel 8 hasil evaluasi menunjukkan bahwa model Gaussian Naive Bayes memperoleh akurasi sebesar 65.53%, model Multinomial Naive Bayes mencapai akurasi 89,65%, dan model Bernoulli Naive Bayes mencatat akurasi sebesar 87,86%.

Tabel 9. Hasil Klasifikasi

Model	Sentimen	Presisi	Recall	F1-score
GaussianNB	Positif	0.91	0.69	0.78
	Negatif	0.13	0.39	0.19
MultinomialNB	Positif	0.90	1.00	0.95
	Negatif	0.00	0.00	0.00
BernoulliNB	Positif	0.90	0.97	0.14
	Negatif	0.26	0.09	0.14

Tabel 9 menunjukkan hasil klasifikasi sentimen menggunakan tiga model Naive Bayes: GaussianNB, MultinomialNB, dan BernoulliNB, dengan metrik presisi, recall, dan F1-score untuk sentimen positif dan negatif.

1. GaussianNB: Kinerja baik pada sentimen positif (presisi 0.91, recall 0.69, F1-score 0.78), namun buruk pada sentimen negatif (presisi 0.13, recall 0.39, F1-score 0.19).

2. MultinomialNB: Sangat baik pada sentimen positif (presisi 0.90, recall 1.00, F1-score 0.95), namun gagal pada sentimen negatif (presisi, recall, F1-score 0.00).
3. BernoulliNB: Cukup baik pada sentimen positif (presisi 0.90, recall 0.97), namun rendah pada F1-score (0.14), dan sangat buruk pada sentimen negatif (presisi 0.26, recall 0.09, F1-score 0.14).

Secara keseluruhan, MultinomialNB unggul untuk sentimen positif, namun semua model kesulitan dalam mengklasifikasikan sentimen negatif.

3.4. Hasil Confusion Matrix

Setelah melakukan tahapan pengujian, tahapan selanjutnya adalah menentukan hasil confusion matrix. Tujuan membandingkan ketiga confusion matrix dari GaussianNB, MultinomialNB, dan BernoulliNB dalam konteks analisis sentiment adalah untuk menentukan model klasifikasi Naïve Bayes mana yang memberikan kinerja terbaik untuk dataset dan representasi fitur yang digunakan.

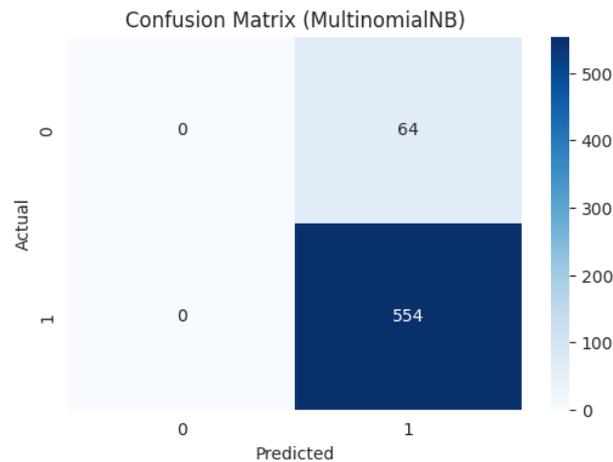
Hasilnya visualisasi pada gambar 5 dalam confusion matrix model Gaussian Naive Bayes (GNB) diimplementasikan untuk klasifikasi sentimen dan dievaluasi menggunakan data uji. Confusion matrix ini memetakan kinerja model dalam mengklasifikasikan sentimen menjadi dua kelas: negatif dan positif, melalui elemen-elemen seperti True Negative (TN), False Positive (FP), False Negative (FN), dan True Positive (TP). Berdasarkan confusion matrix, diperoleh TN sebesar 25, FP sebesar 39, FN sebesar 174, dan TP sebesar 380. Data ini mengungkap adanya ketidakseimbangan kelas (class imbalance) yang signifikan, dengan jumlah data positif (554) jauh melebihi data negatif (64), yang perlu dipertimbangkan dalam interpretasi metrik evaluasi. Hasil pengujian menunjukkan bahwa model Gaussian Naive Bayes memberikan sebesar akurasi 65,53%.

Lalu pada gambar 6 model Multinomial Naive Bayes (MultinomialNB) yang dievaluasi untuk tugas klasifikasi sentimen menunjukkan performa yang sangat buruk dan bias terhadap sentimen positif. Analisis *confusion matrix* mengungkapkan bahwa model gagal total dalam memprediksi sentimen negatif. Secara spesifik, model tidak berhasil memprediksi satupun data negatif dengan benar (*True Negative* atau $TN = 0$), melainkan salah mengklasifikasikan seluruh data negatif sebagai positif (*False Positive* atau $FP = 64$). Sebaliknya, semua data positif diprediksi dengan benar sebagai positif (*True Positive* atau $TP = 554$), dan tidak ada data positif yang salah diprediksi sebagai negatif (*False Negative* atau $FN = 0$). Kondisi ini, dengan TN dan FN bernilai nol, mengindikasikan adanya masalah fundamental pada model, kemungkinan besar disebabkan oleh bias dalam data training, kesalahan dalam proses *preprocessing* data, atau bahkan kesalahan dalam implementasi model itu sendiri. Hasil pengujian menunjukkan bahwa model Multinomial Naïve Bayes memberikan sebesar akurasi 89,65%.

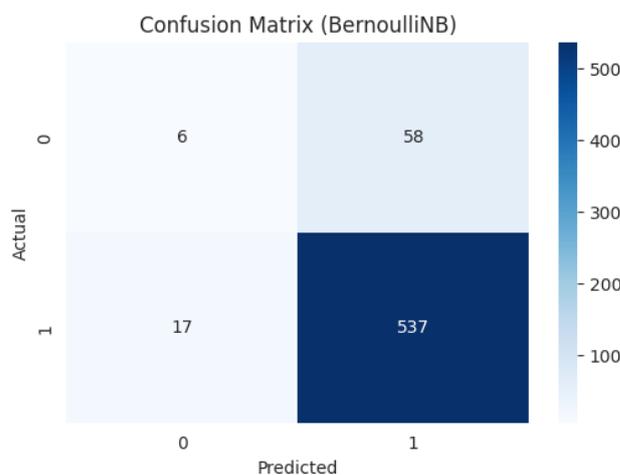
Model Bernoulli Naive Bayes (BernoulliNB) yang digunakan untuk klasifikasi sentimen menunjukkan performa yang cukup baik, meskipun masih terdapat beberapa area yang perlu ditingkatkan. Berdasarkan *confusion matrix* gambar 7 model berhasil memprediksi 6 data negatif dengan benar (*True Negative*, TN) dan 537 data positif dengan benar (*True Positive*, TP). Namun, model juga melakukan kesalahan dengan memprediksi 58 data negatif sebagai positif (*False Positive*, FP) dan 17 data positif sebagai negatif (*False Negative*, FN). Hasil pengujian menunjukkan bahwa model Bernoulli Naïve Bayes memberikan performa terbaik dengan akurasi 87,86%.



Gambar 3. Confusion Matrix GaussianNB



Gambar 4. Confusion Matrix MultinomialNB



Gambar 5. Confusion Matrix BernoulliNB

Berdasarkan evaluasi terhadap ketiga model Naive Bayes yang diterapkan dalam klasifikasi sentimen, masing-masing model menunjukkan tingkat kinerja yang berbeda. Model GaussianNB mencatat akurasi terendah, yaitu 65,53%. Model ini mengalami kesulitan yang signifikan dalam mengklasifikasikan data positif dan negatif dengan akurat, terutama disebabkan oleh ketidakseimbangan kelas yang cukup besar. Sementara itu, model MultinomialNB menunjukkan performa yang jauh lebih baik dengan akurasi mencapai 89,65%. Namun, model ini menghadapi tantangan dalam memprediksi data negatif secara tepat, karena seluruh data negatif dikategorikan sebagai positif, yang mengakibatkan tingkat *False Positive* yang tinggi. Di sisi lain, model BernoulliNB mencatat akurasi sebesar 87,86%, sedikit di bawah MultinomialNB, tetapi tetap menunjukkan kinerja yang memadai. BernoulliNB menunjukkan kemampuan yang lebih seimbang dalam memprediksi data positif dan negatif, meskipun masih terdapat kesalahan dalam mengklasifikasikan data negatif sebagai positif. Secara keseluruhan, model MultinomialNB memberikan hasil terbaik dalam hal akurasi, sementara model BernoulliNB lebih efektif dalam menjaga keseimbangan antara prediksi data positif dan negatif, sehingga menghasilkan klasifikasi sentimen yang lebih dapat diandalkan.

4. KESIMPULAN

Penelitian ini mengkaji sentimen masyarakat terhadap peningkatan gaji guru honorer di Twitter dengan menggunakan algoritma Naive Bayes. Dari total 3091 tweet yang dianalisis, ditemukan bahwa 88,35% mengandung sentimen positif, sementara 11,65% bersifat negatif. Tiga varian model Naive Bayes diuji, yaitu Gaussian, Multinomial, dan Bernoulli. Hasil evaluasi menunjukkan bahwa model Bernoulli Naive Bayes memberikan hasil terbaik dengan akurasi mencapai 87,86%, diikuti oleh Multinomial Naive Bayes yang memiliki akurasi 89,65%. Meskipun Multinomial menunjukkan akurasi yang lebih tinggi, model ini tidak mampu memprediksi sentimen negatif dengan baik, sehingga dianggap kurang dapat diandalkan dibandingkan dengan

Bernoulli yang lebih efektif dalam mengklasifikasikan data negatif. Sementara itu, Gaussian Naïve Bayes menunjukkan performa terendah dengan akurasi 65,53%. Studi ini memberikan informasi berharga bagi para pengambil keputusan untuk memahami pandangan publik mengenai kesejahteraan guru honorer melalui analisis sentimen yang berbasis pada machine learning. Dengan memanfaatkan model yang lebih handal, seperti Bernoulli Naïve Bayes, kebijakan yang lebih tepat dapat dirumuskan untuk meningkatkan kesejahteraan guru honorer. Namun, penelitian ini memiliki batasan karena hanya mengandalkan data dari Twitter. Penelitian mendatang dapat mempertimbangkan untuk mengeksplorasi platform lain seperti Facebook, forum diskusi, atau situs berita guna memperoleh hasil yang lebih representatif dan memperluas pemahaman mengenai sentimen publik terkait isu yang sama. Selain itu, penting untuk memperhatikan aspek keseimbangan kelas dalam model klasifikasi agar prediksi yang dihasilkan lebih akurat dan dapat diandalkan di berbagai kategori sentimen.

DAFTAR PUSTAKA

- [1] M. Muslihudin And S. Hartati, "Guru Honorer Di Kabupaten Pesawaran Menggunakan," *J. Teknol.*, Vol. 9, Pp. 83–88, 2021.
- [2] S. P. Annas Solihin, "Menghitung Harapan: Dampak Kenaikan Gaji Guru Terhadap Masa Depan Pendidikan Indonesia," Unesa. Accessed: Jan. 23, 2025. [Online]. Available: https://S2dikdas.Fip.Unesa.Ac.Id/Post/Menghitung-Harapan-Dampak-Kenaikan-Gaji-Guru-Terhadap-Masa-Depan-Pendidikan-Indonesia?Utm_Source=Chatgpt.Com
- [3] G. A. Fauzan, "Guru Honorer Dalam Lingkaran Ketidakadilan," *J. Educ.*, Vol. 4, No. 1, Pp. 197–208, 2021, Doi: 10.31004/Joe.V4i1.418.
- [4] Admalinea, "Kesejahteraan Guru Honorer Di Indonesia," Lpm Detak Alinea Fisip Unand. Accessed: Jan. 23, 2025. [Online]. Available: https://Detakalinea.Fisip.Unand.Ac.Id/Index.Php/Joomlart-Content/List-All-Categories/19-Berita/82-Kesejahteraan-Guru-Honorer-Di-Indonesia?Utm_Source=Chatgpt.Com
- [5] J. A. Nursiyono And C. Chotimah, "Analisis Sentimen Netizen Twitter Terhadap Pemberitaan Ppn Sembako Dan Jasa Pendidikan Dengan Pendekatan Social Network Analysis Dan Naive Bayes Classifier," *J Stat. J. Ilm. Teor. Dan Apl. Stat.*, Vol. 14, No. 1, Pp. 52–58, 2021, Doi: 10.36456/Jstat.Vol14.No1.A3868.
- [6] O. Kaka, A. Ananda Putra Tanggu Mara, K. Wulla Rato, P. Studi Teknik Informatika, And S. Tinggi Manajemen Informatika Komputer Stella Maris Sumba Email Penulis Korespondensi, "Analisis Sentimen Dampak Perkembangan Teknologi Informasi Dan Komunikasi Terhadap Kemajuan Belajar Siswa Smk Rada Pamba Dengan Metode Naive Baiyes," *J. Ilmu Komput. Dan Sist. Inf.*, Vol. 6, Pp. 191–199, 2023.
- [7] D. R. Andriyani, M. Afdal, And S. Monalisa, "Analisis Sentimen Masyarakat Terhadap Penghapusan Honorer Berdasarkan Opini Dari Twitter Menggunakan Naïve Bayes Classifier," *Build. Informatics, Technol. Sci.*, Vol. 5, No. 1, Pp. 49–58, 2023, Doi: 10.47065/Bits.V5i1.3541.
- [8] M. Taufik Sugandi, Martanto, And U. Hayati, "Analisis Sentimen Komentar Pengguna Youtube Terhadap Kebijakan Baru Badan Penyelenggara Jaminan Kesehatan Sosial Menggunakan Naïve Bayes," *J. Inform. Dan Rekayasa Perangkat Lunak*, Vol. 6, No. 1, Pp. 218–227, 2024.
- [9] A. A. Rizal, G. S. Nugraha, R. A. Putra, And D. P. Anggraeni, "Twitter Sentiment Analysis In Tourism With Polynomial Naïve Bayes Classifier," *Jtim J. Teknol. Inf. Dan Multimed.*, Vol. 5, No. 4, Pp. 343–353, 2024, Doi: 10.35746/Jtim.V5i4.478.
- [10] A. Sitanggang, Y. Umidah, Y. Umidah, R. I. Adam, And R. I. Adam, "Analisis Sentimen Masyarakat Terhadap Program Makan Siang Gratis Pada Media Sosial X Menggunakan Algoritma Naïve Bayes," *J. Inform. Dan Tek. Elektro Terap.*, Vol. 12, No. 3, 2024, Doi: 10.23960/Jitet.V12i3.4902.
- [11] 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.
- [12] Alfandi Safira And F. N. Hasan, "Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier," *Zo. J. Sist. Inf.*, Vol. 5, No. 1, Pp. 59–70, 2023, Doi: 10.31849/Zn.V5i1.12856.
- [13] A. L. Fairuz, R. D. Ramadhani, And N. A. F. Tanjung, "Analisis Sentimen Masyarakat Terhadap Covid-19 Pada Media Sosial Twitter," *J. Dinda Data Sci. Inf. Technol. Data Anal.*, Vol. 1, No. 1, Pp. 42–51, 2021, Doi: 10.20895/Dinda.V1i1.180.
- [14] D. Darwis, N. Siskawati, And Z. Abidin, "Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional," *J. Tekno Kompak*, Vol. 15, No. 1, P. 131, 2021, Doi: 10.33365/Jtk.V15i1.744.

-
- [15] S. W. Ritonga, . Y., M. Fikry, And E. P. Cynthia, “Klasifikasi Sentimen Masyarakat Di Twitter Terhadap Ganjar Pranowo Dengan Metode Naïve Bayes Classifier,” *Build. Informatics, Technol. Sci.*, Vol. 5, No. 1, 2023, Doi: 10.47065/Bits.V5i1.3535.
- [16] R. R. Salam, M. F. Jamil, Y. Ibrahim, R. Rahmaddeni, S. Soni, And H. Herianto, “Analisis Sentimen Terhadap Bantuan Langsung Tunai (Blt) Bahan Bakar Minyak (Bbm) Menggunakan Support Vector Machine,” *Malcom Indones. J. Mach. Learn. Comput. Sci.*, Vol. 3, No. 1, Pp. 27–35, 2023, Doi: 10.57152/Malcom.V3i1.590.
- [17] F. N. Hidayat And S. Sugiyono, “Analisis Sentimen Masyarakat Terhadap Perekrutan Pppk Pada Twitter Dengan Metode Naive Bayes Dan Support Vector Machine,” *J. Sains Dan Teknol.*, Vol. 5, No. 2, Pp. 665–672, 2023, Doi: 10.55338/Saintek.V5i2.1359.
- [18] M. A. S. Putra, I. Permana, M. Mustakim, And M. Afdal, “Analisis Sentimen Masyarakat Mengenai Gerakan Childfree Di Media Sosial X Menggunakan Algoritma Nbc Dan Svm,” *Malcom Indones. J. Mach. Learn. Comput. Sci.*, Vol. 4, No. 4, Pp. 1189–1198, 2024, Doi: 10.57152/Malcom.V4i4.1356.
- [19] N. Krisman And J. Gulo, “Analisis Sentimen Masyarakat Terhadap Kebijakan Kenaikan Uang Kuliah Tunggal (Ukt) Pada Media Sosial Analysis Of Public Sentiment Towards The Policy Of Single Tuition Increase (Ukt) On Social Media X Using The Naive Bayes Classifier Method,” Vol. 3, No. 12, Pp. 714–722, 2024.
- [20] K. A. Lubis, M. T. A. Bangsa, And A. Yudertha, “Analisis Sentimen Opini Masyarakat Terhadap Pindahannya Ibu Kota Indonesia Dengan Menggunakan Klasifikasi Naïve Bayes,” *J. Teknoinfo*, Vol. 18, No. 1, Pp. 226–238, 2024, [Online]. Available: [https://ejurnal.Teknokrat.Ac.Id/Index.Php/Teknoinfo/Index](https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index)
- [21] C. F. Hasri And D. Alita, “Penerapan Metode Naïve Bayes Classifier Dan Support Vector Machine Pada Analisis Sentimen Terhadap Dampak Virus Corona Di Twitter,” *J. Inform. Dan Rekayasa Perangkat Lunak*, Vol. 3, No. 2, Pp. 145–160, 2022, Doi: 10.33365/Jatika.V3i2.2026.