



Analisis Sentimen Platform X Mengenai Pro Kontra Rekrutmen Guru Melalui Marketplace Menggunakan Metode Naïve Bayes

Yusuf Ramadhan Nasution, Aidil Halim Lubis, Tengku Fira Eliza*

Fakultas Sains dan Teknologi, Ilmu Komputer, Universitas Islam Negeri Sumatera Utara, Medan, Indonesia

Email: ¹ramadhannst@uinsu.ac.id, ²aidilhalimlubis@uinsu.ac.id, ^{3,*}tengkufiraeliza58@gmail.com

Email Penulis Korespondensi: tengkufiraeliza58@gmail.com

Abstrak—Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi, Nadiem Anwar Makarim telah membahas terobosan baru terkait sistem rekrutmen guru melalui marketplace. Pada tanggal 24 Mei 2023 sejak pertama kali kebijakan ini disampaikan, sistem marketplace guru menimbulkan pro kontra di tengah masyarakat. Salah satu media sosial yang ramai membicarakan mengenai topik ini adalah media sosial X. Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap opini pengguna Media Sosial X mengenai marketplace guru. Analisis sentimen dilakukan dengan menganalisis 640 data opini. Data tersebut diklasifikasikan menggunakan Metode Naive Bayes. Hasil pengujian menunjukkan pada pengujian dua kelas, yaitu positif dan negatif. Hasil dari klasifikasi dari 640 data opini menunjukkan bahwa sebanyak 340 data diklasifikasikan sebagai data positif dan 270 data sebagai klasifikasi negatif. Hal ini menunjukkan bahwa pengguna Platform X yang memberikan opini lebih pro terhadap kebijakan tersebut. Kemudian, dilakukan pembagian data, yaitu 90% data latih dan 10% data uji. Berdasarkan analisis yang telah dilakukan, maka didapatkan nilai akurasi 74%, precision 80%, recall 35%, dan f-1 score 48%.

Kata Kunci: Marketplace; Guru; Naive Bayes; X; Media Sosial

Abstract—Minister of Education, Culture, Research and Technology, Nadiem Anwar Makarim has discussed new breakthroughs regarding the teacher recruitment system through the marketplace. On May 24, 2024 since this policy was first presented, the teacher marketplace system has raised pros and cons in society. One of the social media that is busy discussing this topic is social media X. This research aims to conduct sentiment analysis of the opinions of Social Media X users regarding the teacher marketplace. Sentiment analysis was carried out by analyzing 640 opinion data. The data is classified using the Naive Bayes Method. The test results show that there are two classes in the test, namely positive and negative. This shows that platform X users who provide opinions are more pro towards this policy. Then, the data is divided, namely 90% training data and 10% test data. Based on the analysis that has been carried out, an accuracy value of 74%, precision 80%, recall 35%, and f-1 score 48% were obtained.

Keywords: Marketplace; Teacher; Naive Bayes; X; Social Media

1. PENDAHULUAN

Penggunaan media sosial merupakan kegiatan umum yang dilakukan banyak orang atau bisa disebut dengan gaya hidup yang sedang trend, khususnya di Indonesia. Chriss Garrett mendefinisikan media sosial sebagai kumpulan alat, layanan, dan interaksi yang dirancang khusus untuk memungkinkan hubungan antara individu yang memiliki minat yang sama [1]. Tidak diragukan lagi, media sosial berfungsi sebagai alat utama yang digunakan individu untuk mempertahankan aktivitas mereka dan berinteraksi dengan orang lain.

Platform media sosial X, yang mengalami perubahan nama pada 22 Juli 2023, berfungsi sebagai media untuk mengekspresikan sentimen publik. Twitter melakukan proses rebranding dengan mengubah namanya menjadi X dan mengganti lambang burung biru konvensional dengan X putih di latar belakang hitam. Ini merupakan komponen dari inisiatif rebranding yang lebih luas, yang dilakukan sebagai respons terhadap pilihan Elon Musk untuk menjadi pemilik platform jejaring sosial ini. Platform media sosial X merupakan media penghubung yang sangat diminati oleh individu di seluruh dunia. Menurut hasil survei We Are Social yang dilakukan pada tahun 2023, penggunaan media sosial secara global di Indonesia telah mencapai 167 juta individu atau setara dengan jumlah penduduk sebesar 276,4 juta jiwa. Sementara itu, survei pada Januari menyebutkan jumlah pengguna X di dunia 556 juta sedangkan Indonesia menjadi peringkat keenam sebagai pengguna X terbanyak di dunia dengan 24 juta pengguna. Sementara itu, X juga menjadi media sosial urutan keenam yang paling banyak digunakan di Indonesia. Rata-rata, X menghasilkan 383 juta tweet setiap hari, dengan basis pengguna melebihi 29,5 juta di Indonesia. Lebih jauh, X menempati peringkat keenam sebagai platform jejaring sosial yang paling sering digunakan di Indonesia. Basis pengguna yang besar dan aksesibilitas yang mudah ke X akan menghasilkan akumulasi sejumlah besar data opini, yang selanjutnya dapat menjadi peluang berharga untuk studi kasus penelitian.

Penelitian ini berfokus pada analisis manfaat dan kekurangan perekrutan guru melalui marketplace, sebuah inisiatif yang baru-baru ini diperkenalkan oleh Nadiem Anwar Makarim, Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi, pada 24 Mei 2023. Marketplace daring adalah platform digital tempat pedagang dapat menawarkan produk mereka tanpa mengharuskan pelanggan memiliki situs web atau toko daring sendiri [2].

Nadiem Makarim telah mengusulkan inovasi baru dalam perekrutan guru dan bermaksud untuk melaksanakannya dengan menggunakan paradigma pasar guru. Nadiem berpendapat bahwa pembentukan pasar guru akan mengatasi tantangan yang dihadapi oleh guru honorer, termasuk ketidakmampuan untuk secara langsung merekrut guru baru karena harus bergantung pada perekrutan guru ASN yang terpusat. Meskipun demikian, usulan untuk merekrut instruktur melalui sistem pasar guru telah menimbulkan keuntungan dan kerugian di seluruh masyarakat. Beberapa orang menganggap pembicaraan tersebut sebagai hal baru dan menyatakan kekhawatiran bahwa hal itu dapat membanjiri instruktur tambahan. Lebih jauh, beberapa orang menganggap pasar guru tidak sesuai dengan martabat guru sebagai panggilan mulia

dan menyatakan kekhawatiran tentang efek buruk sistem ini terhadap mutu dan kesejahteraan guru. Sebaliknya, ada juga yang percaya bahwa sistem ini akan memberi insentif kepada instruktur untuk meningkatkan kompetensi profesional mereka agar menarik minat sekolah yang mempertimbangkan layanan mereka.

Terkait dengan pokok bahasan yang dibahas, pengguna X menggunakannya untuk mengomunikasikan informasi dalam bentuk kritik atau rekomendasi yang membangun terhadap kebijakan yang dikomunikasikan oleh Nadiem Makarim. Oleh karena itu, penelitian ini menawarkan pendekatan yang diusulkan untuk mengkaji sudut pandang pengguna X terhadap kelebihan dan kekurangan rekrutmen guru melalui marketplace dengan menggunakan analisis sentimen. Postingan para pengguna X dapat diklasifikasikan dengan menggunakan metode analisis sentimen yang secara otomatis mengidentifikasi, mengekstraksi, dan memproses informasi tekstual untuk menemukan informasi emosional dalam ekspresi pikiran [3]. Dianggap lebih mudah dan sederhana dalam mengekspresikan pandangan mereka, X juga dianggap lebih populer di kalangan masyarakat Indonesia. Oleh karena itu sangat efisien jika menggunakan X sebagai media untuk mengambil data mengenai pro kontra rekrutmen guru melalui marketplace [4]. Proses analisis sentimen adalah pendekatan metodis yang digunakan untuk mendeteksi perasaan dan mengklasifikasikan polaritas teks dalam dokumen atau kalimat untuk membuat kelompok terpisah yang mewakili sentimen positif dan negatif. Saat ini, analisis sentimen banyak digunakan oleh para peneliti sebagai bidang studi utama dalam bidang ilmu komputer [5]. Analisis sentimen adalah penerapan teknik otomatis untuk memahami, mengekstrak, dan menganalisis data tekstual yang diberikan untuk mendapatkan informasi sentimen yang melekat dalam sebuah kalimat. Analisis Sentimen digunakan untuk menemukan informasi berharga yang dibutuhkan dari data yang tidak terstruktur, sehingga diharapkan pada penelitian ini dapat diketahui sentimen pengguna twitter terhadap aplikasi Ruang Guru [6].

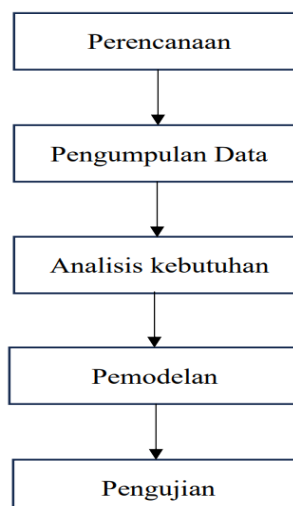
Penelitian yang dilakukan oleh Dwi Normawati dan Surya Allit Prayogi pada tahun 2021, penelitian ini bertujuan untuk membuat sistem yang dapat mengolah data dengan menerapkan analisis sentimen. Penelitian ini juga menggunakan Metode Naive Bayes untuk menganalisis tingkat sentimen terhadap data yang dilakukan dengan teknik crawling pada Platform X. Penelitian ini mendapatkan akurasi sebesar 82%, presisi 92%, dan recall 52%. Kemudian penelitian yang dilakukan oleh Nila Hardi dan kawan-kawan pada tahun 2021, penelitian ini dilakukan untuk menemukan dan mengkategorikan opini negatif atau positif serta metode yang digunakan adalah Metode Naive Bayes. Nilai akurasi yang diperoleh dari penelitian ini sebesar 50,26%.

Dari kajian penelitian terdahulu inilah penulis kemudian mengambil penelitian mengenai analisis sentimen dengan menggunakan Metode Naive Bayes. Pendekatan ini tergolong pembelajaran terbimbing karena melibatkan seorang pengawas yang bertindak sebagai pengajar selama proses pembelajaran. Selain itu, algoritma Naive Bayes menawarkan kecepatan klasifikasi yang unggul, sehingga meningkatkan efektivitas sistem analisis sentimen. Menurut Olson Delen (2008) menjelaskan Naive Bayes untuk setiap kelas keputusan, menghitung probabilitas dengan syarat bahwa kelas keputusan adalah benar, mengingat vektor informasi obyek. Algoritma ini mengasumsikan bahwa atribut obyek adalah independen. Sedangkan jumlah frekuensi yang terdapat pada tabel keputusan merupakan faktor penentu dari probabilitas terhadap perkiraan akhir [7]. Selanjutnya penelitian ini diharapkan mendapatkan nilai akurasi yang tinggi dengan melakukan analisis sentimen pada X.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dilakukan dengan tahapan yang tersaji pada gambar 1, terdapat 5 proses tahapan secara umum dalam penelitian ini yaitu; perencanaan, pengumpulan data, analisis kebutuhan, pemodelan, dan pengujian.



Gambar 1. Tahapan Penelitian



a. Perencanaan

Sebelum membuat model, langkah penelitian diawali dengan adanya suatu kebijakan dan perencanaan untuk membuat model. Perencanaan dianggap penting karena akan menjadi penentu dan sekaligus memberi arah terhadap tujuan yang ingin dicapai. Dengan demikian suatu kerja akan berantakan dan tidak terarah jika tidak ada perencanaan yang matang, perencanaan yang matang dan disusun dengan baik akan memberi pengaruh terhadap ketercapaian tujuan [8]. Tahapan-tahapan dalam melakukan perencanaan, yaitu menentukan topik penelitian, menentukan objek penelitian, rumusan masalah, judul penelitian, tujuan, dan batasan masalah.

b. Pengumpulan Data

Proses pengumpulan data melibatkan dua metodologi berbeda, yaitu studi literatur dan penelusuran data. Studi literatur adalah referensi ilmiah yang digunakan untuk memperoleh pengetahuan tertentu. Sebelum memulai penelitian, biasanya dilakukan tinjauan literatur untuk meneliti teori-teori mendasar [9]. crawling adalah teknik otomatis untuk mengumpulkan informasi dari web, di mana data yang terkumpul dipengaruhi oleh perihal apa yang pengguna cari. Alat ini biasanya dikenal dengan nama crawler, suatu sistem yang dirancang dengan suatu program khusus. Pengumpulan data untuk studi ini terdiri dari data yang diperoleh dari server media sosial X, termasuk pengguna, kiriman, dan fitur deskriptif lainnya [10].

c. Analisis Kebutuhan

Analisis kebutuhan dilakukan untuk memastikan solusi optimal untuk mengatasi masalah dan memenuhi persyaratan yang ditetapkan [11].

d. Pemodelan

Pemodelan sistem adalah proses membangun model suatu sistem. Pemodelan adalah representasi atau deskripsi komprehensif yang menjelaskan suatu objek, sistem, atau konsep melalui penyederhanaan atau idealisasi [12].

e. Pengujian

Pengujian model dilakukan untuk menilai kemandirian kinerja model. Proses pengujian model melibatkan evaluasi tingkat presisi yang ditunjukkan oleh pendekatan tersebut.

2.2 Metode Analisis Data

Tujuan dari penelitian ini adalah untuk menganalisis data opini pengguna Platform X mengenai pro dan kontra rekrutmen guru melalui marketplace dengan menggunakan Metode Klasifikasi Naive Bayes. Algoritma Naive Bayes mencari nilai dengan probabilitas tertinggi agar dapat mengelompokkan data uji dengan kategori yang tepat [13]. Metode NBC menempuh dua tahap dalam proses klasifikasi teks, yaitu tahap pelatihan dan tahap klasifikasi. Dalam pengembangan basis data, Naive Bayes melibatkan pembelajaran yang diawasi jenis pembelajaran mesin [14]. Pada tahap pelatihan dilakukan proses analisis terhadap sampel dokumen berupa pemilihan vocabulary, yaitu kata yang mungkin muncul dalam koleksi dokumen sampel yang sedapat mungkin dapat menjadi representasi dokumen [15]. Metode Naive Bayes merupakan teknik klasifikasi probabilistik yang secara langsung menghitung serangkaian probabilitas dengan menggabungkan frekuensi dan kombinasi nilai dari data empiris yang relevan [16]. Keuntungan mendasar dari metode ini terletak pada kapasitasnya untuk menghitung estimasi parameter dengan lebih sedikit data pelatihan dengan menggabungkan informasi dari sampel dengan data yang sudah ada sebelumnya [17]. Persamaan teori Bayes disajikan di sini [18]. Menghitung nilai prior probability adalah sebagai berikut:

$$P(X) = \frac{dx}{d} \quad (1)$$

Keterangan untuk rumus di atas adalah $P(X)$ sebagai prior probability, X sebagai hipotesis data merupakan suatu kategori (kelas), dx sebagai jumlah dengan kategori (kelas) x , d sebagai jumlah seluruh dataset
Menghitung nilai conditional probability dengan laplacian smoothing:

$$P(\text{Term } H_n|X) = \frac{H_n|X + 1}{(c) + |V|} \quad (2)$$

Keterangan untuk rumus di atas adalah $P(\text{Term } H_n|X)$ sebagai conditional probability kata ke- n dengan kelas x , $(H_n|X)$ sebagai frekuensi kemunculan kata pada kelas x , (c) sebagai bobot (IDF) dari kata pada kelas x , dan $|V|$ sebagai nilai IDF dari seluruh kata yang terdapat pada data. Menghitung nilai posterior probability:

$$P(\text{Term } H|X) = P(X) \times P(\text{Term } H_1|X) \times P(\text{Term } H_2|X) \quad (3)$$

Keterangan untuk rumus di atas adalah $P(\text{Term } H|X)$ sebagai posterior probability, $P(X)$ sebagai prior probability, dan $P(\text{Term } H_n|X)$ sebagai conditional probability. Persamaan-persamaan di atas adalah model dari teorema Naive Bayes Multinomial yang akan digunakan dalam proses klasifikasi.

2.3 Metode Pengujian

Confusion Matrix adalah matriks yang digunakan untuk mengevaluasi proses model klasifikasi dengan mengukur jumlah pengamatan data uji yang akurat dan tidak akurat. Berdasarkan matriks ini, seseorang dapat memastikan tingkat keunggulan dalam kinerja model klasifikasi. Matriks kebingungan adalah representasi tabular yang mengukur proporsi data uji yang diidentifikasi dengan benar dan proporsi data uji yang diklasifikasikan secara salah dari himpunan data

tertentu [19]. Evaluasi klasifikasi pada penelitian ini menggunakan metode Confusion Matrix. Cara kerja metode ini yaitu matriks dari prediksi akan dibandingkan dengan kelas asli yang berisi informasi nyata dan prediksi nilai klasifikasi. Setelah sistem berhasil melakukan klasifikasi tweet, dibutuhkan ukuran untuk menentukan seberapa benar atau tepat klasifikasi yang telah dibuat oleh sistem [20].

Tujuan utama dari matriks kebingungan ini adalah untuk membandingkan hasil klasifikasi yang dihasilkan oleh model yang digunakan dengan hasil klasifikasi yang tepat. Matriks kebingungan berisi istilah-istilah yang merangkum hasil dari proses kategorisasi. Keempat istilah ini adalah True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN). Matriks kebingungan menjelaskan mekanisme operasional model dalam proses prediksi. Matriks kebingungan berfungsi tidak hanya untuk memasukkan kesalahan yang dihasilkan oleh model, tetapi juga mengidentifikasi karakteristik khusus dari kesalahan ini. Confusion matrix pada dasarnya informasi akurasi yang dihasilkan didapatkan dari perbandingan hasil klasifikasi yang dihasilkan oleh sistem dengan hasil yang seharusnya [21]. Perhitungan presisi, recall, akurasi, dan skor f-1 dapat dilakukan menggunakan matriks kebingungan yang disediakan dalam tabel 1. Presisi adalah ukuran ketepatan antara respons sistem dengan informasi yang diinginkan oleh pengguna. Recall adalah ukuran keberhasilan sistem dalam menemukan kembali informasi yang relevan. Sedangkan akurasi adalah tingkat kesesuaian antara nilai prediksi dan nilai aktual [22].

Tabel 1. Confusion Matrix

Nilai Prediksi	Nilai Aktual	
	Positive	Negative
Positive	True Positive	False Positive
Negative	False Negative	True Negative

Tabel di atas merupakan tabel confusion matrix yang terdiri dari nilai prediksi dan nilai aktual. Berikut adalah keterangan dari tabel tersebut:

- True Positive (TP): Jumlah data yang bernilai Positif dan diprediksi benar sebagai Positif.
- False Positive (FP): Jumlah data yang bernilai negatif tetapi diprediksi sebagai positif.
- False Negative (FN): Jumlah data yang bernilai positif tetapi diprediksi sebagai negatif.
- True Negative (TN): Jumlah data yang bernilai negatif dan diprediksi benar sebagai negatif.

3. HASIL DAN PEMBAHASAN

3.1 Data Crawling

Pengumpulan statistik dimulai pada tanggal 29 Juni 2023. Total data yang diperoleh selama prosedur penelusuran data berjumlah 1877 titik data. Kemudian, data tersebut diproses kembali untuk membuang data duplikat serta data yang tidak relevan dengan topik penelitian. Dengan demikian, jumlah total data adalah 640. Kata kunci yang digunakan dalam prosedur perayapan data meliputi "teacher marketplace", "guru", "nadiem", dan "kemendikbud". Berikut ini adalah ilustrasi posting yang dibuat oleh pengguna platform X, yang dirinci dalam tabel 2.

Tabel 2. Contoh Hasil Data Crawling posting X

Data	Posting
1	OTAK BISNIS BUSUK NADIEM DITERAPKAN DALAM RANAH PENDIDIKAN.....?????? GURU DIPERJUAL BELIKAN LAYAKNYA BARANG DAGANGAN"???? Marketplace merendahkan harkat martabat guru????? https://t.co/rpU2mrdJAq "
2	@mahlid_ emang kadang nadiem bener bikin guru masuk marketplace kalau begini ????
3	@byyy_me Kesalahan paling awal dari program ini... Menteri Nadiem menamainya dengan... MARKETPLACE
4	Program Marketplace Guru Inisiasi Nadiem Dikritik, Dinilai Hanya Solusi Jangka Pendek https://t.co/tn9J2wM8tl
5	Marketplace Guru, Urgensikah di Tengah Rendahnya Kualitas Pendidikan? https://t.co/Nb88HY90qZ https://t.co/xJqipjxcZY

3.2 Text Preprocessing

Selama fase praproses, data teks yang diperoleh dari hasil pengikisan X akan mengalami pemrosesan dan transformasi ke dalam format yang lebih ramping dan terorganisir. Hal ini bertujuan untuk memperbaiki kualitas data dan memudahkan proses analisis sentimen. Tugas dalam praproses teks meliputi pelipatan huruf besar/kecil, pembersihan, tokenisasi, normalisasi, penghapusan kata henti, dan stemming.

a. Case Folding dan Cleaning

Pelipatan huruf besar-kecil bertujuan untuk mengubah semua huruf pada halaman menjadi huruf kecil, sedangkan pembersihan secara khusus menargetkan penghapusan kata-kata yang tidak relevan seperti referensi, URL, tagar, dan emotikon. Hasil dari operasi pelipatan huruf besar-kecil dan pembersihan disajikan dalam Tabel 3 dari tinjauan ini.

**Tabel 3.** Hasil Case Folding dan Cleaning

Data	Hasil Case Folding dan Cleaning
1	otak bisnis busuk nadiem diterapkan dalam ranah pendidikan guru diperjual belikan layaknya barang dagangan marketplace merendahkan harkat martabat guru
2	emang kadang nadiem bener bikin guru masuk marketplace kalau begini
3	kesalahan paling awal dari program ini menteri nadiem menamanya dengan marketplace
4	program marketplace guru inisiasi nadiem dikritik dinilai hanya solusi jangka pendek
5	marketplace guru urgensikah tengah rendahnya kualitas pendidikan

b. Tokenizing

Tokenizing adalah proses seleksi pemotongan kata dalam kalimat. Tabel 4 adalah hasil dari proses tokenizing yang sebelumnya sudah melewati proses case folding.

Tabel 4. Hasil Proses Tokenizing

Data	Hasil Tokenizing
1	['otak', 'bisnis', 'busuk', 'nadiem', 'diterapkan', 'dalam', 'ranah', 'pendidikan', 'guru', 'diperjual', 'belikan', 'layaknya', 'barang', 'dagangan', 'marketplace', 'merendahkan', 'harkat', 'martabat', 'guru']
2	['emang', 'kadang', 'nadiem', 'bener', 'bikin', 'guru', 'masuk', 'marketplace', 'kalau', 'begini']
3	['kesalahan', 'paling', 'awal', 'dari', 'program', 'ini', 'menteri', 'nadiem', 'menamanya', 'dengan', 'marketplace']
4	['program', 'marketplace', 'guru', 'inisiasi', 'nadiem', 'dikritik', 'dinilai', 'hanya', 'solusi', 'jangka', 'pendek']
5	['marketplace', 'guru', 'urgensikah', 'tengah', 'rendahnya', 'kualitas', 'pendidikan']

c. Normalization

Tujuan dari normalisasi adalah untuk memperbaiki istilah-istilah yang dianggap tidak baku dan tidak sesuai dengan ketentuan Kamus Besar Bahasa Indonesia. Tabel 5 menampilkan hasil dari prosedur normalisasi data.

Tabel 5. Hasil Proses Normalization

Data	Hasil Normalization
1	['otak', 'bisnis', 'busuk', 'nadiem', 'diterapkan', 'dalam', 'ranah', 'pendidikan', 'guru', 'diperjual', 'belikan', 'layaknya', 'barang', 'dagangan', 'marketplace', 'merendahkan', 'harkat', 'martabat', 'guru']
2	['memang', 'kadang', 'nadiem', 'benar', 'bikin', 'guru', 'masuk', 'marketplace', 'kalau', 'begini']
3	['kesalahan', 'paling', 'awal', 'dari', 'program', 'ini', 'menteri', 'nadiem', 'menamanya', 'dengan', 'marketplace']
4	['program', 'marketplace', 'guru', 'inisiasi', 'nadiem', 'dikritik', 'dinilai', 'hanya', 'solusi', 'jangka', 'pendek']
5	['marketplace', 'guru', 'urgensikah', 'tengah', 'rendahnya', 'kualitas', 'pendidikan']

d. Stopword Removal

Kata-kata yang dihapus merujuk pada penghapusan yang disengaja atas kata-kata yang berlebihan atau tidak penting. Tabel 6 menyajikan hasil dari prosedur penghapusan kata-kata tidak penting.

Tabel 6. Hasil Proses Stopword Removal

Data	Hasil Stopword Removal
1	['otak', 'bisnis', 'busuk', 'nadiem', 'diterapkan', 'ranah', 'pendidikan', 'guru', 'diperjual', 'belikan', 'layaknya', 'barang', 'dagangan', 'marketplace', 'merendahkan', 'harkat', 'martabat', 'guru']
2	['memang', 'kadang', 'nadiem', 'benar', 'bikin', 'guru', 'masuk', 'marketplace', 'kalau', 'begini']
3	['kesalahan', 'paling', 'awal', 'program', 'menteri', 'nadiem', 'menamanya', 'marketplace']
4	['program', 'marketplace', 'guru', 'inisiasi', 'nadiem', 'dikritik', 'dinilai', 'solusi', 'jangka', 'pendek']
5	['marketplace', 'guru', 'urgensikah', 'tengah', 'rendahnya', 'kualitas', 'pendidikan']

e. Stemming

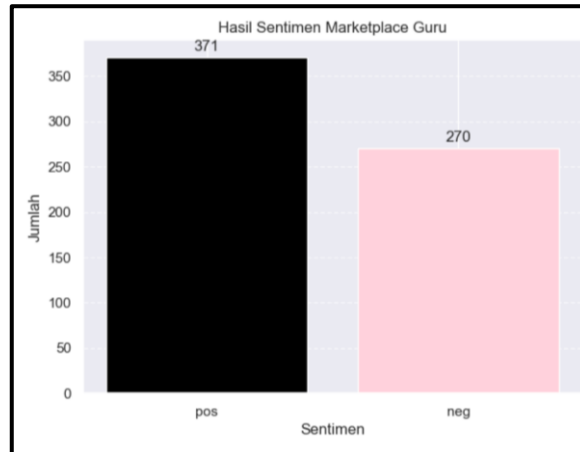
Stemming juga bisa diartikan sebagai proses menghilangkan infleksi kata ke bentuk dasarnya (Isnain et al., 2021). Tabel 7 memberikan hasil prosedur stemming.

Tabel 7. Hasil Proses Stemming

Data	Hasil Stemming
1	otak bisnis busuk nadiem terap ranah didik guru jual belikan layak barang dagang marketplace rendah harkat martabat guru
2	memang kadang nadiem benar bikin guru masuk marketplace kalau begini
3	salah paling awal program menteri nadiem nama marketplace
4	program marketplace guru inisiasi nadiem kritik nilai solusi jangka pendek
5	marketplace guru urgensi tengah rendah kualitas didik

3.3 Labeling

Data hasil text preprocessing selanjutnya memasuki tahap pelabelan yang dilakukan menggunakan library TextBlob. Labeling data juga digunakan untuk mengetahui sifat dari teks yang sudah dikumpulkan dan sentimen yang digunakan adalah sentimen yang bernilai positif atau negatif. TextBlob merupakan salah satu pelabelan yang dilakukan secara otomatis yang digunakan untuk memproses data yang bersifat tekstual. Label positif diberikan kepada postingan yang pro terhadap kebijakan mengenai rekrutmen guru melalui marketplace, sedangkan label negatif diberikan kepada postingan yang kontra terhadap kebijakan mengenai rekrutmen guru melalui marketplace. Berikut ini adalah visualisasi data untuk jumlah sentimen positif dan negatif dalam bentuk plot-bar.



Gambar 2. Plot-bar Sentimen Marketplace Guru

3.4 TF-IDF

Selama fase ini, kata-kata individual diberi bobot. Kumpulan data berisi 2492 istilah. Terminologi Pembobotan Frekuensi-Frekuensi Dokumen Terbalik (TF-IDF) adalah teknik yang digunakan untuk mengubah data tekstual menjadi data numerik dengan tujuan memberikan bobot pada kata-kata atau fitur individual dalam kumpulan data tertentu.

3.6 Klasifikasi Naive Bayes

Selanjutnya, penelitian ini akan dilanjutkan dengan pelaksanaan klasifikasi menggunakan metode pembelajaran mesin Naive Bayes. Pendekatan klasifikasi yang digunakan dalam penelitian ini adalah Naive Bayes Multinomial. Pendekatan ini menunjukkan tingkat presisi yang terpuji dan melibatkan perhitungan yang mudah. Secara khusus, data dibagi menjadi dua set berbeda, yang dikenal sebagai data pelatihan dan data pengujian. Data pelatihan digunakan untuk pelatihan berulang dan pengoptimalan model, sedangkan data pengujian digunakan untuk tujuan kategorisasi menurut model. Menggunakan rasio data pelatihan terhadap data pengujian sebesar 9:10. Sebanyak 90% data ditujukan untuk tujuan pelatihan, sedangkan 10% sisanya disisihkan untuk pengujian. Ringkasan data pelatihan dan pengujian disediakan dalam Tabel 8.

Tabel 8. Jumlah Data Latih dan Data Uji

Data Latih	Data Uji
576	65

Setelah melakukan split dataset, data latih yang telah diberikan label akan dilatih menggunakan pemodelan Naive Bayes. Hasil akhir dari pemodelan ini menghasilkan tingkat akurasi sebesar 74%.

3.7 Confusion Matrix

Langkah terakhir yang dilakukan dalam pekerjaan ini adalah mengevaluasi pemodelan dengan menggunakan matriks kebingungan. Hasil matriks kebingungan mengacu pada hasil data uji yang diproyeksikan. Hasil yang diperoleh direpresentasikan sebagai matriks 2x2 yang mencakup kategori sentimen positif dan negatif. Selanjutnya, akurasi, recall, presisi, dan skor f1 akan dihitung berdasarkan matriks kebingungan yang dihasilkan. Matriks 2x2 yang dihasilkan ditampilkan dalam tabel 9:

Tabel 9. Hasil Klasifikasi Confusion Matrix

Nilai Prediksi	Nilai Aktual	
	Positif	Negatif
Positif	8	2
Negatif	15	40



Dari matriks tersebut, maka diperoleh hasil accuracy, precision, recall, dan f-1 score. Hasil tersebut disajikan dalam tabel 10 berikut ini.

Tabel 10. Hasil Perhitungan Confusion Matrix

	Precision	Recall	F1-score	Support
Negative	0.80	0.35	0.48	23
Positive	0.73	0.95	0.82	42
Accuracy			0.74	65
Macro avg	0.76	0.65	0.65	65
Weighted avg	0.75	0.74	0.70	65

Hasil akhir dari confusion matrix memperoleh nilai presisi sebesar 80%, recall sebesar 35%, F1-score sebesar 48%, dan akurasi yang diperoleh sebesar 74%.

4. KESIMPULAN

Penelitian saat ini menggunakan Metode Naive Bayes untuk menganalisis sikap pro dan kontra tentang perekrutan guru melalui pasar. Metode ini menggunakan teknik klasifikasi dengan memberikan label positif dan negatif serta menerapkan probabilitas bersyarat. Hasil labelisasi berbentuk sentimen positif dan negatif. Penerapan pendekatan pelabelan ini menghasilkan 371 contoh sentimen baik dalam kelas data dan 279 contoh sentimen negatif. Selanjutnya, hasil pengujian menghasilkan akurasi 74%, presisi 80%, recall 35%, dan statistik f1 48%. Oleh karena itu, dapat dipastikan bahwa pendekatan klasifikasi Naive Bayes sangat relevan meskipun kinerja akurasi tidak sempurna. Salah satu kemungkinan jalan untuk studi lebih lanjut adalah membangun sistem yang terdiri dari antarmuka pengguna berbasis web, desktop, atau seluler.

REFERENCES

- [1] N. A. Pahlevi, Pengaruh Media Sosial dan Gerakan Massa terhadap Gerakan Hakim. Jakarta: Cipta Media Nusantara, 2021.
- [2] A. Mubarak, M. D. Dinantara, Susanto, Zulfitra, and Sahroni, "Strategi Pemasaran Melalui Marketplace dalam Upaya Peningkatan Penjualan pada UMKM Desa Pabuaran," *Abdi Laksana*, vol. 2, no. 3, pp. 534–539, 2021.
- [3] 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.
- [4] R. Puspita and A. Widodo, "Perbandingan Metode KNN, Decision Tree, dan Naive Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS," *J. Inform. Univ. Pamulang*, vol. 5, no. 4, p. 646, 2021, doi: 10.32493/informatika.v5i4.7622.
- [5] U. Verwardina, F. Edi, and R. Watianthos, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naive Bayes," vol. 5, pp. 157–163, 2021, doi: 10.30865/mib.v5i1.2604.
- [6] A. P. Giovani, A. Ardiansyah, T. Haryanti, L. Kurniawati, and W. Gata, "Analisis Sentimen Aplikasi Ruang Guru Di Twitter Menggunakan Algoritma Klasifikasi," *J. Teknoinfo*, vol. 14, no. 2, p. 115, 2020, doi: 10.33365/jti.v14i2.679.
- [7] J. Sihombing, "Klasifikasi Data Antropometri Individu Menggunakan Algoritma Naive Bayes Classifier," *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 2, no. 1, pp. 1–10, 2021, doi: 10.37148/bios.v2i1.15.
- [8] D. Darwisah, K. Imron Rosadi, and H. Ali, "Berfikir Kesisteman Dalam Perencanaan Dan Pengembangan Pendidikan Islam," *J. Manaj. Pendidik. Dan Ilmu Sos.*, vol. 2, no. 1, pp. 225–237, 2021, doi: 10.38035/jmpis.v2i1.444.
- [9] U. Y. Sundari et al., *Metodologi Penelitian*. Padang: CV. Gita Lentera, 2024.
- [10] R. Danang et al., "Implementasi Web Crawling Pada Website Bps Jawa Timur Pada Data Infasi Di Jawa Timur," *UMAT J. Pengabdian Al-Shobar Untuk Masy.*, vol. 1, no. 2, pp. 6–13, 2024, [Online]. Available: <https://journal.alshobar.or.id/index.php/umat>
- [11] I. K. D. G. Supartha et al., *Buku Ajar Analisa Perancangan Sistem*. Jambi: PT. Sonpedia Publishing Indonesia, 2023.
- [12] N. Ayunandita and S. Dadi Riskiono, "Permodelan Sistem Informasi Akademik Menggunakan Extreme Programming Pada Madrasah Aliyah (Ma) Mambaul Ulum Tanggamus," *J. Inform. dan Rekayasa Perangkat Lunak*, vol. 2, no. 2, pp. 196–204, 2021, [Online]. Available: <http://jim.teknokrat.ac.id/index.php/informatika>
- [13] B. Ramadhani and R. R. Suryono, "Komparasi Algoritma Naive Bayes dan Logistic Regression Untuk Analisis Sentimen Metaverse," *J. Media Inform. Budidarma*, vol. 8, no. 2, p. 714, 2024, doi: 10.30865/mib.v8i2.7458.
- [14] M. K. Insan, U. Hayati, and O. Nurdian, "Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di," *J. Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 478–483, 2023.
- [15] D. Duei Putri, G. F. Nama, and W. E. Sulistiono, "Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier," *J. Inform. dan Tek. Elektro Terap.*, vol. 10, no. 1, pp. 34–40, 2022, doi: 10.23960/jitet.v10i1.2262.
- [16] F. D. Aulya and R. A. Putri, "Penerapan Metode Naive Bayes Dan Forward Chaining Untuk Diagnosis Penyakit Gangguan Ibu Hamil Berbasis Web," *INTECOMS J. Inf. Technol. Comput. Sci.*, vol. 6, no. 2, pp. 650–656, 2023, [Online]. Available: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=FP0rNUcAAAAJ&pagesize=100&citation_for_view=FP0rNUcAAAAJ:hMod-77fHWUC
- [17] N. Puspitasari, H. Hamdani, H. Hatta, A. Septiarini, and S. Sumaini, "Penerapan Metode Teorema Bayes Untuk Mendeteksi Hama Pada Tanaman Padi Mayas Kalimantan Timur," *SINTECH (Science Inf. Technol. J.)*, vol. 4, no. 2, pp. 155–162, 2021, doi: 10.31598/sintechjournal.v4i2.919.
- [18] Y. S. T. Allo, V. Sofica, N. Hasan, and M. Septiani, "Penggunaan Metode Naive Bayes Dalam Mengklasifikasi Pengangguran Pada Desa Bojong Kulur," *Bianglala Inform.*, vol. 10, no. 1, pp. 30–35, 2022, doi: 10.31294/bi.v10i1.12333.



- [19] 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.
- [20] N. Hendrastuty, A. Rahman Isnain, and A. Yanti Rahmadhani, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *J. Inform. J. Pengemb. IT*, vol. 6, no. 3, pp. 150–155, 2021.
- [21] A. Damuri, U. Riyanto, H. Rusdianto, and M. Aminudin, "Implementasi Data Mining dengan Algoritma Naïve Bayes Untuk Klasifikasi Kelayakan Penerima Bantuan Sembako," *JURIKOM (Jurnal Ris. Komputer)*, vol. 8, no. 6, p. 219, 2021, doi: 10.30865/jurikom.v8i6.3655.
- [22] N. Hidayah and Dodiman, "Implementasi Algoritma Multinomial Naïve Bayes, TF-IDF dan Confusion Matrix dalam Pengklasifikasian Saran Monitoring dan Evaluasi Mahasiswa Terhadap Dosen Teknik Informatika Universitas Dayanu Ikhsanuddin," *J. Akad. Pendidik. Mat.*, vol. 10, no. 1, pp. 8–15, 2024.