

ANALISIS SENTIMEN MASYARAKAT DI PLATFORM X TERHADAP PENGGUNAAN BANSOS UNTUK MEMENANGKAN SALAH SATU CAPRES TERTENTU DI PILPRES 2024 MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER

Fikri Alim Adiyatma, Syariful Alam, Mutiara Andayani Komara

Teknik Informatika, Sekolah Tinggi Teknologi Wastukencana Purwakarta
Jalan Cikopak No.53, Mulyamekar, Kec.Babakancikao, Kab.Purwakarta, Jawa Barat, Indonesia
fikriadiyatma15@gmail.com

ABSTRAK

Media sosial, khususnya *platform Twitter*, telah menjadi *platform* utama bagi pengguna untuk mengungkapkan pendapat mereka mengikuti tren berita secara global. Di Indonesia, penggunaan *Twitter* meluas, mencerminkan signifikansi platform ini dalam mempengaruhi opini publik. Salah satu isu yang hangat dibahas adalah penggunaan bantuan sosial (BANSOS) dalam konteks politik, terutama terkait Pilpres 2024. Tujuan dari penelitian ini adalah untuk menganalisis sentimen masyarakat di *platform Twitter* terhadap penggunaan BANSOS untuk memenangkan salah satu calon presiden menggunakan metode *Naïve Bayes Classifier*. Metode ini digunakan untuk mengklasifikasikan sentimen positif, negatif, dan netral dari 722 dokumen yang telah melalui tahap *preprocessing*, termasuk *cleaning*, *labelling*, *tokenizing*, dan *stemming*. Hasil analisis menunjukkan bahwa sebagian besar pendapat yang diekspresikan cenderung positif, dengan 71,3% *tweet* mengungkapkan pandangan positif terhadap strategi penggunaan BANSOS dalam kampanye politik. Evaluasi menggunakan *Confusion Matrix* menunjukkan tingkat akurasi sebesar 80,82%, presisi 80,28%, *recall* 100%, dan *f1-score* 89,06%. Hasil ini mengindikasikan bahwa mayoritas pengguna *Twitter* memiliki pandangan positif terhadap penggunaan BANSOS untuk mendukung calon presiden di Pilpres 2024. Temuan ini menawarkan wawasan yang lebih mendalam mengenai dinamika opini publik di media sosial serta hubungannya dengan proses politik di Indonesia.

Kata kunci : *twitter, bansos, presiden, naïve bayes classifier*

1. PENDAHULUAN

Media sosial saat ini berkembang pesat, salah satunya *platform X*, yang lebih dikenal sebagai *Twitter*. *Platform X* adalah media sosial di mana orang dari seluruh dunia dapat mengungkapkan pendapat mereka. Data yang dihasilkan di *Twitter* sangat berguna jika dianalisis melalui *data mining*, karena bisa diubah menjadi informasi penting. Di *Twitter*, pendapat tentang berita atau tren apa pun dapat diamati dengan baik. Indonesia menempati peringkat kelima terbesar di dunia dalam hal pengguna *Twitter*, setelah Amerika Serikat, Brasil, Jepang, dan Inggris, dengan 19,5 juta pengguna dari keseluruhan 300 juta pengguna di seluruh dunia [1].

Analisis sentimen adalah sebuah kegiatan yang dilakukan dengan komputasi dengan pendekatan secara statistik atau pendekatan secara leksikal untuk melakukan analisis emosi, sentimen, atau opini dari seseorang yang umumnya berupa teks. Hasil dari sentimen analisis memiliki tiga klasifikasi, yaitu: Positif, Negatif, dan Netral [2]

Pemilihan presiden merupakan momen penting dalam kehidupan demokrasi sebuah negara, dimana masyarakat berhak memilih pemimpin yang akan mewakili kepentingan dan aspirasi mereka. Dalam konteks PILPRES 2024, penggunaan bantuan sosial (BANSOS) dijadikan alat politik untuk memenangkan salah satu calon presiden telah menjadi perbincangan hangat di kalangan masyarakat. Menurut UU Perpres Nomor 63 tahun 2017, bantuan sosial merupakan

bantuan dalam bentuk uang, barang, atau dukungan langsung kepada individu, keluarga, kelompok atau masyarakat miskin, tidak mampu atau rentan terhadap risiko sosial.

Penggunaan BANSOS sebagai alat politik menimbulkan beragam pandangan di masyarakat, terutama di kalangan pengguna *Twitter*. Ada yang mendukungnya sebagai strategi mendapatkan dukungan politik, sementara yang lain menentanginya sebagai manipulasi merugikan kepentingan publik. Dalam konteks PILPRES 2024, penting memahami tanggapan masyarakat terhadap penggunaan BANSOS dalam kampanye politik. Pandangan terhadap isu ini di *Twitter* menunjukkan peran penting media sosial dalam membentuk opini publik. Analisis sentimen terhadap penggunaan BANSOS di *Twitter* relevan untuk memahami dinamika politik menjelang pemilihan presiden. Menggunakan metode *naïve bayes*, Penelitian ini bertujuan untuk menganalisis sentimen masyarakat di *Twitter* terkait penggunaan BANSOS menjelang PILPRES 2024 dengan menggunakan algoritma *Naive Bayes Classifier*. Penggunaan metode ini dalam analisis sentimen media sosial seperti *Twitter* bertujuan untuk memahami pola opini publik terkait penggunaan BANSOS dalam politik. *Naive Bayes Classifier* merupakan salah satu metode yang efektif untuk mengklasifikasikan teks ke dalam kategori sentimen positif, negatif, atau netral. Melalui pendekatan ini, penelitian ini akan mengungkap respons dan pemahaman masyarakat

mengenai peran BANSOS dalam konteks politik serta bagaimana persepsi ini dapat mempengaruhi dinamika politik di Indonesia.

Pada penelitian mengenai analisis sentimen dilakukan oleh [3] mempelajari analisis sentimen yang diterapkan untuk objek vaksinasi Astra Zeneca pada Twitter Menggunakan Metode Naïve Bayes dan K-NN. Studi tersebut menunjukkan bahwa pengujian algoritma *Naïve Bayes Classifier* menunjukkan akurasi sebesar 81%, sedangkan algoritma K-NN menunjukkan akurasi sebesar 74,78%

Berdasarkan hasil penelitian sebelumnya, algoritma Naive Bayes direkomendasikan untuk digunakan. Penelitian ini menerapkan algoritma Naive Bayes untuk menganalisis sentimen terkait isu BANSOS pada PILPRES 2024.

2. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen merupakan studi komputasional tentang opini, sentimen, dan emosi yang diekspresikan dalam teks. Tugas utama dalam analisis sentimen adalah mengkategorikan polaritas teks dalam dokumen, kalimat, atau pendapat. Analisis ini digunakan untuk memberikan gambaran umum dari media sosial guna mengetahui apakah perasaan yang diungkapkan cenderung bersifat positif atau negatif. [4].

Analisis sentimen melibatkan mengekspresikan pendapat melalui penyampaian opini, sikap, evaluasi, emosi, sentimen, subjektivitas, dan pandangan dalam sebuah teks. [5].

2.2. Bantuan Sosial

Bantuan sosial adalah bantuan yang diberikan dalam bentuk uang atau barang dari pemerintah daerah kepada individu, keluarga, kelompok, dan/atau masyarakat. Bantuan ini bersifat tidak kontinu dan selektif, dengan tujuan melindungi mereka dari potensi risiko sosial. [6].

2.3. Platform X

Twitter adalah *platform* media sosial yang memungkinkan pengguna dari seluruh dunia untuk menyampaikan pandangan mereka. Data yang dihasilkan di *Twitter* sangat berharga jika dianalisis, karena dapat diubah menjadi informasi penting melalui analisis opini. Pendapat mengenai berita, peluncuran produk, atau tren tertentu dapat diamati secara efektif di *Twitter* [7].

Twitter, juga dikenal sebagai X, adalah situs *web* yang dimiliki dan dioperasikan oleh *Twitter Inc*, yang didirikan pada tahun 2006 oleh Jack Dorsey. Berbasis di San Bruno, California, dekat San Francisco, situs ini pertama kali diluncurkan. *Microblogging* melalui *Twitter* telah menjadi salah satu media sosial yang populer di kalangan pengguna di Indonesia. *Twitter* adalah *platform* yang menyediakan layanan *microblogging online*, memungkinkan pengguna untuk membagikan konten dengan panjang hingga 280

karakter. Jejaring sosial ini sangat mudah digunakan, karena informasi dapat dibagikan dengan cepat dan menyebar luas dalam waktu singkat. [8]

2.4. Naïve Bayes Classifier

Naïve Bayes Classifier adalah algoritma yang sangat populer dalam *data mining* karena kemudahannya dalam penggunaan dan kecepatan pemrosesan yang tinggi, membuatnya sederhana dan sangat efisien. Metode *Naïve Bayes* didasarkan pada konsep probabilitas yang diperkenalkan oleh British Thomas Bayes, yang memungkinkan untuk memperkirakan masa depan berdasarkan pengalaman masa lalu. [9]

Naïve Bayes Classifier adalah salah satu metode klasifikasi statistik yang dapat memprediksi probabilitas suatu data termasuk dalam kelas tertentu berdasarkan perhitungan probabilitas. [10]

$$P(X|H) = \frac{P(X|H) \times P(H)}{P(X)} \quad (1)$$

Keterangan:

X = Data Sample dengan kelas (label) yang tidak diketahui.

H = Bahwa X adalah data dengan kelas (label) C.

P(H|X) = Bahwa hipotesa benar (valid) untuk data sampel X yang diamati.

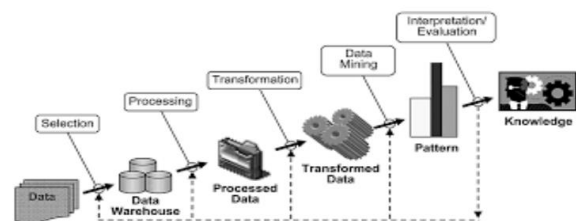
P(X|H) = Peluang data sampel X, bila diasumsikan bahwa hipotesa benar (valid).

P(H) = Peluang dari hipotesa H.

P(X) = Peluang data sampel yang diamati.

2.5. Knowledge Discovery In Databases

Knowledge Discovery in Database (KDD) adalah proses yang melibatkan hasil dari *data mining*, yaitu mengekstraksi pola data. Proses ini kemudian mengubah hasil tersebut menjadi informasi yang mudah dipahami secara akurat [11]



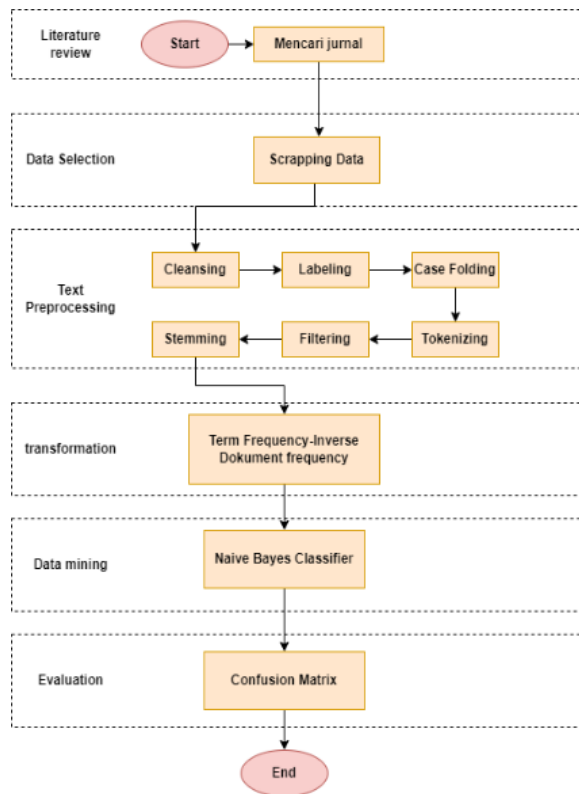
Gambar 1. Tahapan kdd

2.6. Evaluasi Confussion Matrix

Confusion Matrix adalah metode yang digunakan untuk menghitung akurasi dalam *Data Mining*. Evaluasi menggunakan *Confusion Matrix* menghasilkan nilai akurasi, presisi, dan *recall*. Akurasi dalam klasifikasi *Data Mining* adalah persentase keakuratan data yang telah diklasifikasikan dengan benar dan diuji hasil klasifikasinya. Presisi, atau *confidence*, adalah proporsi kasus yang diprediksi positif dan benar-benar positif. *Recall*, atau *sensitivity*, adalah proporsi kasus positif yang berhasil diprediksi dengan benar. [12]

3. METODE PENELITIAN

Dalam penelitian ini, metode yang digunakan adalah Knowledge Discovery in Database (KDD). Metode ini berfokus pada penemuan informasi baru dari data besar yang belum dikenal sebelumnya. Penelitian ini mencakup rencana penelitian yang terdiri dari serangkaian aktivitas untuk mencapai tujuan penelitian. Proses yang diterapkan melibatkan langkah-langkah berikut:



3.1. Scrapping Data

Dataset yang digunakan dalam penelitian ini terdiri dari data yang diperoleh melalui proses pengumpulan. Data tersebut dikumpulkan dari komentar atau opini pengguna media sosial X, dengan menggunakan *Twitter API* untuk mengakses informasi tersebut. Setelah opini dikumpulkan, data disimpan dalam format *file CSV*.

3.2. Text Preprocessing

Text Preprocessing yang bertujuan untuk mempersiapkan teks menjadi data yang terstruktur dan dapat diproses pada tahapan *Cleansing*, *Labeling*, *Case Folding*, *Stemming*, *Filtering*, *Tokenizing*.

3.3. Tranformation (TF-IDF)

Selanjutnya dilakukan proses *term weighting*. Proses ini dilakukan untuk menilai *term* atau bobot kata dari sebuah dokumen terhadap seluruh dokumen pada korpus. Proses ini dilakukan menggunakan bahasa pemrograman *python*.

3.4. Data Mining

Tahap klasifikasi ini melakukan terhadap dokumen yang mengandung kategori tertentu, kata-kata dalam dokumen tersebut yang dapat dilakukan pengklasifikasian, sekumpulan dari kata pada dokumen tersebut dapat digunakan untuk menentukan kategori dikarenakan memiliki nilai makna tertentu. Untuk menentukan sentimen ini dari dokumen *tweet* yaitu menggunakan metode *naïve bayes*.

3.5. Evaluation

Evaluasi bertujuan untuk mengecek kebenaran hasil klasifikasi. Pada tahap ini hasil evaluasi menggunakan metode *cross validation* yang nantinya akan diukur dengan menggunakan *Confusion Matrix* untuk mengetahui hasil akurasi.

4. HASIL DAN PEMBAHASAN

4.1. Scrapping Data

Data dikumpulkan melalui proses pengumpulan informasi di media sosial *Twitter* menggunakan *Twitter API (Application Programming Interface)*. Proses ini dilakukan dengan menggunakan kata kunci "BANSOS". Pengambilan data dilakukan pada tanggal 10 Maret 2024.

4.2. Preprocessing Text

Text Preprocessing yang bertujuan untuk mempersiapkan teks menjadi data yang terstruktur dan dapat diproses pada tahapan berikutnya:

4.2.1. Cleansing

Cleansing merupakan Penghapusan kata duplikat dan penghapusan kata yang tidak sesuai. *Cleansing* data dilakukan dengan using pemrograman *Python*. Berikut contoh data yang belum melalui tahapan *Cleansing* dan contoh data yang sudah melalui tahapan *Cleansing*.

Tabel 1. Proses *cleansing*

Before	After
@J_D3pp Dana habis buat bagi2 BANSOS pemilu ... □	Dana habis buat bagi BANSOS pemilu
@AbdulRachimmks Mungkin kedunguan menular atau serangan BANSOS menjelang pemilu begitu kencang ?	Mungkin kedunguan menular atau serangan BANSOS menjelang pemilu begitu kencang
@Andria75777 Kebodohan di piara negara untuk di ambil suaranya di tukar dng BANSOS,kasus,jabatan.sa at pemilu tiba.	Kebodohan di piara negara untuk di ambil suaranya di tukar dng BANSOS,kasus,jabatan.sa at pemilu tiba
@PakarINTELEk @jokowi mungkin kecapaian lagi sebelum pemilu bagi BANSOS?	jokowi mungkin kecapaian lagi sebelum pemilu bagi BANSOS

4.2.2. Labeling

Pada tahap *Labeling* ini atau pelabelan pada setiap dokumen *tweet* di kategorikan dengan hasil sentiment menjadi positif dan negatif.

Table 2. Proses *Labeling*

Dokumen	Label
Pemilu 2024 akan dikenang dengan dua kata Medsos dan berjuta juta BANSOS	Negatif
Syukur syukur dapet BANSOS, dari pada gak dapet sama sekali	Negatif
Bodoh rasanya kalau BANSOS bisa membuat orang menang pemilu.	Positif
tanggung jawab... beras mahal krn cadangan beras dipake BANSOS pemilu Zholim,	Positif

4.2.3. Case Folding

Tahap *Case Folding* adalah Mengubah seluruh huruf dari “a” sampai dengan “z” dalam dokumen menjadi huruf kecil.

Table 3. Hasil *case folding*

Dokumen
pemilu 2024 akan dikenang dengan dua kata medsos dan berjuta juta bansos
syukur syukur dapet bansos, dari pada gak dapet sama sekali
bodoh rasanya kalau bansos bisa membuat orang menang pemilu.
tanggung jawab beras mahal krn cadangan beras dipake bansos pemilu zholim,

4.2.4. Tokenizing

Tokenizing adalah proses membagi teks menjadi bagian-bagian kecil yang disebut token. Token ini biasanya berupa kata atau frasa singkat yang bisa dianalisis lebih lanjut. Contoh dari proses tokenisasi dapat dilihat pada tabel berikut.

Table 4. Proses tokenizing

pemilu	dikenang	kata	medsos	dan	berjuta
bansos					
syukur	dapet	bansos	pada	gak	dapet
sama	sekali				
bodoh	rasanya	kalau	bansos	bisa	buat
orang	menang	pemilu			

4.2.5. Filtering

Tahap *Filter* adalah tahap mengambil kata-kata penting dari hasil *Tokenizing* menggunakan algoritma *stopword removal* (membuang kata yang kurang penting). *Stopword* adalah kata-kata yang tidak deskriptif yang dapat dibuang dalam pendekatan *bag-of-word*.

Table 5. Hasil filtering

pemilu	dikenang	medsos	berjuta	bansos
syukur	dapat	bansos	dapet	sekali
bodoh	rasanya	kalau	bansos	
orang	menang	pemilu		

4.2.6. Stemming

Tahap *Stemming* adalah tahap mencari root kata dari tiap kata hasil *Filter*. Pada tahap ini dilakukan proses pengembalian berbagai bentukan kata kedalam suatu representasi yang sama

Table 6. Proses stemming

Before	After
pemilu 2024 akan dikenang dengan dua kata medsos dan berjuta juta bansos	pemilu 2024 kenang dengan dua medsos juta bansos
syukur syukur dapet bansos, dari pada gak dapet sama sekali	Syukur dapet bansos, dapet sama sekali
bodoh rasanya kalau bansos bisa membuat orang menang pemilu.	bodoh rasa kalau bansos membuat orang menang pemilu.
tanggung jawab beras mahal krn cadangan beras dipake bansos pemilu zholim,	tanggung jawab beras mahal cadang beras pake bansos pemilu zholim

4.3. Transformation (TF-IDF)

Setelah tahap *preprocessing* selesai, langkah berikutnya adalah tahap pembobotan kata. Dalam tahap ini, setiap kata akan diberikan bobot (nilai). Proses ini dimulai dengan menghitung *term frequency*, kemudian *inverse document frequency*, dan terakhir menggabungkan keduanya dalam *term frequency-inverse document frequency (TF-IDF)*.

1. Tf- IDF

Setelah tahap *preprocessing* selesai, langkah berikutnya adalah tahap pembobotan kata. Dalam tahap ini, setiap kata akan diberikan bobot (nilai). Proses ini dimulai dengan menghitung *term frequency*, kemudian *inverse document frequency*, dan terakhir menggabungkan keduanya dalam *term frequency-inverse document frequency (TF-IDF)*.

Table 7. Sampel dokumen *tweet*

Dokumen	Label
rakyat sudah lapar bos nunggu 2 atau 3 tahun lagi ndasmu makanya masyarakat bodoh yg mau coblos no 2 bagi2 BANSOS terahir ditipu besar mahal ah ah ah strategi jokowi sangat buruk sekali dan apakah masih mau bela no 2 makzulkan jokowi dan pemilu buruk terburuk di alam ini	Negatif
ahy sby sudah tinggalkan rakyat yg medambakan keadilan ahy sby sekarang dlm lindungan pabrik mobil smk bodong konstitusi bodong langgar etika pencapresan 02 BANSOS 496 trilyun dibagi jelang pemilu ngaku gak curang harapan besar kod allah s	positif
pemilu dah beres bos beda ceritanya kalo pas masih pemilu tp bagi bagi BANSOS	positif
ambang batas presiden di hapus juga amp wewenang presiden soal anggaran di mutilasi agar pemilu tidak ada kecurangan presiden buat bagi2 BANSOS saat pemilu	Postif

Pada tabel 7 merupakan sample dokumen dimana sample dokumen ini selanjutnya akan di cari bobot dalam kata yang sering muncul berikut tabel 8 hasil *term frequency*.

Table 8. Hasil trem frequency

Tweet	TF			
	D1	D2	D3	D4
Rakyat	1	1	0	0
Pemilu	0	0	1	2
Lapar	1	0	0	0
Bansos	1	1	1	1
Bodong	1	0	0	0
Presiden	0	0	0	1
Jokowi	2	0	0	0

Pada tabel 9 didapatkan bobot pada kata yang sering muncul pada kalimat. Kemudian tahap selanjutnya mencari idf. setelah semua kata diberikan bobot, langkah selanjutnya adalah menghitung *Document Frequency (DF)* atau seberapa sering kata tersebut muncul dalam sebuah dokumen. Selanjutnya, total data dari seluruh dokumen dibagi dengan *Document Frequency (DF)* untuk mendapatkan nilai dokumen (DF). Detailnya dapat dilihat pada tabel berikut:

Table 9. Proses hasil idf

Tweet	TF				DF	D/DF	IDF(Log D/DF)	1+ DF
	D1	D2	D3	D4				
rakyat	1	1	0	0	2	2	0,301	1,301
pemilu	0	0	1	2	3	1,333	0,477	1,4077
lapar	1	0	0	0	1	4	0	1
bansos	1	1	1	1	4	1	0,602	1,602
bodong	1	0	0	0	1	4	0	1
presiden	0	0	0	1	1	4	0	1
jokowi	2	0	0	0	2	2	0,301	1,301

Setelah di ketahui hasil idf kemudian menghitung hasil tf idf. Tahap *Term Frequency Inverse Document Frequency (TF-IDF)* bertujuan untuk memperoleh nilai tf-idf dengan mengalikan bobot dari setiap dokumen dengan nilai idf dari kata tersebut. Hasil perhitungan ini dapat dilihat pada tabel di bawah ini.

Table 10. term frequency inverse document frequency

Tweet	W			
	D1	D2	D3	D4
Rakyat	0,301	0,301	0	0,602
Pemilu	0	0	0,01	0,477
Lapar	0,301	0,301	0,301	1,204
BANSOS	0,301	0	0	0,301
Bodong	0,301	0	0	1,204
Presiden	0	0	0	1,204
Jokowi	0,602	0	0	0,602

4.4. Data Mining (Klasifikasi Naïve Bayes Classifier)

Dokumen-dokumen yang berisi opini dari media sosial X yang diperoleh melalui wawancara kemudian diklasifikasikan ke dalam kategori yang relevan berdasarkan istilah-istilah yang ada di dalam dokumen tersebut. Dalam melakukan klasifikasi ini, penulis menggunakan algoritma *Naïve Bayes* untuk membuat penilaian berdasarkan data yang telah dikumpulkan sebelumnya.

Sebelum melakukan klasifikasi data, kita terlebih dahulu membagi dataset dibagi menjadi dua bagian: data pelatihan dan data pengujian. Dimana prosesi split data menjadi 90% data latih dan 10% data uji atau sebanyak 649 data untuk latih, 73 data untuk uji. Hasil model klasifikasi dapat dilihat pada gambar dibawah :

```

+ Inisialisasi model dan pelatihan
CMB = ComplementNB()
CMB.fit(x_train_tfidf, y_train)

+ Prediksi
predictedCMB = CMB.predict(x_test_dtm)
accuracyCMB = accuracy_score(predictedCMB, y_test)

+ Print akurasi
print('complementNB model accuracy is', str('{:04.2f}'.format(accuracyCMB*100)) + '%')
print('-----')

complementNB model accuracy is 80.82%

```

Gambar 2. Hasil klasifikasi

Hasil akurasi yang didapat adalah 80,82%. Ini menunjukkan bahwa penggunaan *algoritma naïve bayes* cukup efektif, karena dari hasil perhitungan tersebut, akurasi terbukti tinggi.

4.5. Evaluation

Proses evaluasi model bertujuan untuk menilai kinerja model klasifikasi yang telah dikembangkan dan mengukur tingkat akurasi. Dalam penelitian ini, penulis memanfaatkan metode *Confusion Matrix* untuk menghitung nilai akurasi, presisi, dan *recall*. Berikut ini adalah ilustrasi dari *Confusion Matrix* yang digunakan.

```

Confusion Matrix:
  0   1
0  2  14
1  0  57

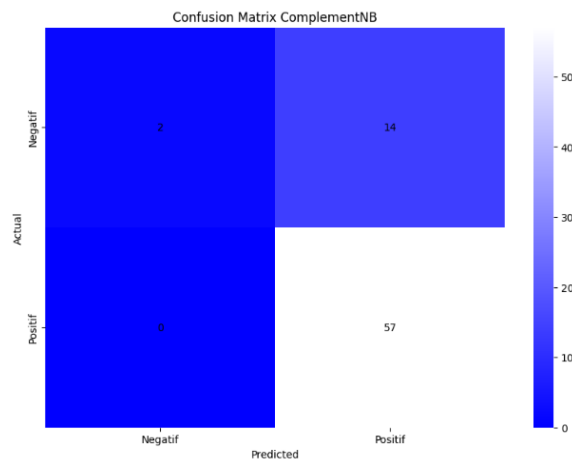
-----
Classification Report:
              precision    recall  f1-score   support

   Negatif         1.00      0.12      0.22         16
   Positif         0.80      1.00      0.89         57

 accuracy                   0.81         73
 macro avg              0.90      0.56      0.56         73
 weighted avg           0.85      0.81      0.74         73

```

Gambar 3. Hasil confusion matrix

Gambar 4. Hasil *confusion matrix*

Kemudian proses perhitungan manual *confusion matrix* di bawah ini:

a. Akurasi

$$Akurasi = \frac{2 + 57}{73} = 80,82\%$$

b. Presisi

$$a. Precision (Positif) = \frac{TP}{TP + FP} = \frac{57}{57 + 14} = 80,28\%$$

$$b. Precision (Negatif) = \frac{TN}{TN + FN} = \frac{2}{2 + 0} = 100,00\%$$

c. Recall

$$a. Recall (Positif) = \frac{TP}{TP + fp} = \frac{57}{57 + 0} = 100,00\%$$

$$b. Recall (Negatif) = \frac{TN}{TN + fN} = \frac{2}{2 + 14} = 12,50\%$$

d. F1-Score

$$F1(Positif) = \frac{2 \times Precision(positif) \times Recall(positif)}{Precision(positif) + Recall(positif)} = \frac{2(80,28 \times 100,00)}{80,28 + 100,00} = 89,06\%$$

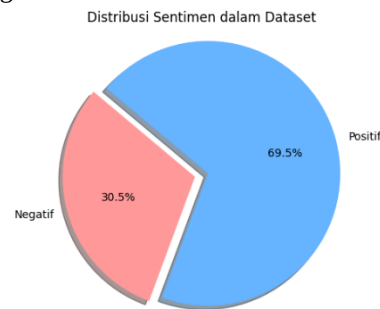
$$F1(Negatif) = \frac{2 \times Precision(Negatif) \times Recall(Negatif)}{Precision(Negatif) + Recall(Negatif)} = \frac{2(100,00 \times 12,50)}{100,00 + 12,50} = 22,22\%$$

Hasil perhitungan manual *confusion matrix* di atas menunjukkan akurasi sebesar 80,82%, *precision* sebesar 80,28%, *recall* sebesar 100%, dan *f1-score* sebesar 89,06%. Hasil ini konsisten dengan hasil yang diperoleh dari hasil perhitungan *python* didalam *google collab*.

4.6 Visualisasi

Visualisasi merupakan sebuah rekayasa gambar, diagram maupun animasi dalam menampilkan sebuah informasi. Untuk penelitian kali ini visualisasi menggunakan *Word Cloud* dan *Diagram*

a. Kategori setimen



Gambar 5. Kategori sentiment

b. Wordcloud

Pada gambar di bawah merupakan kata yang sering muncul pada sentiment positif.



Gambar 6. Wordcloud positif

Kemudian kata yang sering muncul pada sentiment negatif dapat dilihat pada gambar di bawah ini.



Gambar 7. Wordcloud negative

5. KESIMPULAN DAN SARAN

Penelitian ini menganalisis sentimen pada dokumen yang membahas opini mengenai penggunaan BANSOS untuk memenangkan salah satu calon presiden (CAPRES) di PILPRES 2024. Data yang dianalisis berjumlah 722 dan telah melewati tahap preprocessing, yang meliputi pembersihan (*cleansing*), pelabelan (*labelling*), pengubahan huruf (*case folding*), tokenisasi (*tokenizing*), penyaringan (*filtering*), dan *stemming*. Analisis menggunakan algoritma *Naïve Bayes Classification* dan evaluasi dilakukan dengan *confusion matrix*. Hasil penelitian menunjukkan

bahwa terdapat 515 data positif (71,3%) dan 207 data negatif (28,77%). Tingkat akurasi mencapai 80,82%, dengan *precision* sebesar 80,28%, *recall* 100%, dan *F1-score* 89,06%. Dari hasil ini, dapat disimpulkan bahwa mayoritas pengguna platform X memiliki opini positif terhadap isu penggunaan BANSOS untuk memenangkan calon presiden di PILPRES 2024. Dukungan ini menunjukkan bahwa mayoritas pengguna di platform X mendukung opini mengenai penggunaan BANSOS untuk tujuan tersebut. Untuk penelitian di masa mendatang yang membahas isu serupa, pertimbangkan menggunakan algoritma pembelajaran mesin yang lebih canggih seperti Random Forest, Support Vector Machines (SVM), atau algoritma deep learning untuk membandingkan kinerjanya dengan Naïve Bayes Classification. Selain itu, perbanyak jumlah data yang dianalisis untuk memperoleh hasil yang lebih representatif dan akurat, karena data yang lebih besar dapat memberikan wawasan lebih mendalam dan mengurangi bias. Kombinasikan juga data dari platform media sosial dengan sumber lain seperti artikel berita, blog, atau forum diskusi untuk mendapatkan perspektif yang lebih menyeluruh.

DAFTAR PUSTAKA

- [1] N. Wahyu, Ela dan W. Wibowo, "Analisis Sentimen Pengguna Twitter terhadap Program Kartu Prakerja di Tengah Pandemi Covid-19 Menggunakan Metode Naïve Bayes Classifier," JURNAL SAINS DAN SENI ITS, Surabaya, 2022.
- [2] F. S. Yerzi dan Y. Sibaroni, "ANALISIS SENTIMEN TERHADAP KEBIJAKAN PEMERINTAH DALAM MENANGANI COVID-19 DENGAN PENDEKATAN LEXICON BASED," e-Proceeding of Engineering, Bandung, 2021.
- [3] S. H. Ramadhani dan M. I. Wahyudin, "Analisis Sentimen Terhadap Vaksinasi Astra Zeneca pada Twitter Menggunakan Metode Naïve Bayes dan K-NN," *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 6, no. 4, hal. 526–534, 2022, doi: 10.35870/jtik.v6i4.530.
- [4] A. Nurian, M. S. Ma'arif, I. N. Amalia, dan C. Rozikin, "ANALISIS SENTIMEN PENGGUNA APLIKASI SHOPEE PADA SITUS GOOGLE PLAY MENGGUNAKAN NAIVE BAYES CLASSIFIER," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 1, Jan 2024, doi: 10.23960/jitet.v12i1.3631.
- [5] R. Metivianis, D. E. Ratnawati, dan B. Rahayudi, "Analisis Sentimen pengguna Twitter terhadap Vaksinasi Sinovac dan AstraZeneca menggunakan Algoritma CART," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Malang, 2022. [Daring]. Tersedia pada: <http://j-ptiik.ub.ac.id>
- [6] I. Winarta, Tagel, Wayan, A. Raka, Gede, Agung, dan I. Sumada, Made, "Evaluasi Kebijakan Penyaluran Dana Bantuan Sosial di Pemerintah Kabupaten Gianyar," *Public Inspiration: Jurnal Administrasi Publik*, Denpasar, Bali, 2020. [Daring]. Tersedia pada: <https://ejournal.warmadewa.ac.id/index.php/public-inspiration>
- [7] A. Malik Zuhdi, E. Utami, dan S. Raharjo, "ANALISIS SENTIMENT TWITTER TERHADAP CAPRES INDONESIA 2019 DENGAN METODE K-NN," *Jurnal INFORMATIKA Politeknik Indonusa Surakarta, Yogyakarta*, 2019. [Daring]. Tersedia pada: <https://doi.org/10.46808/informa.v5i2.73>
- [8] S. Rezeki, irwansyah, Reza, Y. Restiviani, dan R. Zahara, "PENGUNAAN SOSIAL MEDIA TWITTER DALAM KOMUNIKASI ORGANISASI (Studi Kasus Pemerintah Provinsi DKI Jakarta Dalam Penanganan Covid-19)," *J. Islam. LAW Stud.*, vol. 20, hal. 63–78, 2020.
- [9] S. W. Ritonga, . Y., M. Fikry, dan E. P. Cynthia, "Klasifikasi Sentimen Masyarakat di Twitter terhadap Ganjar Pranowo dengan Metode Naïve Bayes Classifier," *Build. Informatics, Technol. Sci.*, vol. 5, no. 1, Jun 2023, doi: 10.47065/bits.v5i1.3535.
- [10] A. Yoseva Simanjuntak dan I. Septian Salomo Simatupang, "IMPLEMENTASI DATA MINING MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER UNTUK DATA KENAIKAN PANGKAT DINAS KETENAGAKERJAAN KOTA MEDAN," *Journal of Science and Social Research*, Purwakarta, 2022. [Daring]. Tersedia pada: <http://jurnal.goretanpena.com/index.php/JSSR>
- [11] Y. B. Utomo, I. Kurniasari, dan I. Yanuartanti, "PENERAPAN KNOWLEDGE DISCOVERY IN DATABASE UNTUK ANALISA TINGKAT KECELAKAAN LALU LINTAS," *J. Tek. Inform. Kaputama*, vol. 7, no. 1, Jan 2023.
- [12] V. Ginting, Saputra, Kusrini, dan Taufiq, "IMPLEMENTASI ALGORITMA C4.5 UNTUK MEMPREDIKSI KETERLAMBATAN PEMBAYARAN SUMBANGAN PEMBANGUNAN PENDIDIKAN SEKOLAH MENGGUNAKAN PYTHON," *J. Teknol. Inf. dan Komun.*, 2020.