e-ISSN : 2722-4368

Penerapan Naïve Bayes Classifier dalam Klasifikasi Sentimen Publik di Twitter terhadap Puan Maharani

Rizki Hidayat[1], Muhammad Fikry[2], Yusra[3*], Febi Yanto[4], Eka Pandu Cynthia[5]
[12345]Universitas Islam Negeri Sultan Syarif Kasim Riau, Jl. HR. Soebrantas No.Km.15, RW. 15, Simpang Baru, Pekanbaru, Indonesia.
[1]11950115210@students.uin-suska.ac.id, [2]muhammad.fikry@uin-suska.ac.id, [3*]yusra@uin-suska.ac.id,
[4] febiyanto@uin-suska.ac.id, [5] eka.pandu.cynthia@uin-suska.ac.id

Abstrak. *Twitter adalah salah satu jejaring sosial terpopuler di Indonesia, dengan 18,45 juta pengguna aktif pada tahun 2022. Politisi berpengaruh Puan Maharani menjadi topik hangat di pesta ulang tahunnya di tengah protes harga bahan bakar. Analisis sentimen dapat membantu memahami keseluruhan sentimen yang diungkapkan di Twitter tentang Puan Maharani. Dua jenis dataset yang digunakan dalam penelitian ini, yaitu dataset tidak seimbang (9000 tweet: 7800 positif, 1200 negatif) dan dataset seimbang (2400 tweet: 1200 positif, 1200 negatif). Metode Naive Bayes classifier digunakan untuk klasifikasi sentimen, meliputi pengumpulan data, pelabelan, preprocessing, pembobotan TF-IDF, seleksi fitur, pembagian data, klasifikasi Naive Bayes, dan evaluasi dengan confusion matrix. Data dibagi dengan rasio 70:30, 80:20 dan 90:10 untuk data latih serta data uji. Feature selection menggunakan threshold 0,001. Merujuk hasil penelitian yang dilaksanakan, bisa disimpulkan bahwsanya analisis sentimen dapat menjadi alat yang efektif untuk memahami pendapat masyarakat khususnya netizen di platform Twitter terkait dengan persepsi terhadap Puan Maharani. Nilai akurasi tertinggi dari dataset tidak seimbang didapatkan yaitu sebesar 88.89% pada rasio pembagian data latih dan data uji 90:10 serta akurasi tertinggi dari dataset seimbang sebesar 81.0% pada rasio pembagian data 90:10.*

Kata Kunci: Puan Maharani; Klasifikasi Sentimen; Naive Bayes Classifier; Twitter.

Abstract. *Twitter is one of the most popular social networks in Indonesia, with 18.45 million active users by 2022. Influential politician Puan Maharani became a hot topic at her birthday party amidst fuel price protests. Sentiment analysis can help understand the overall sentiment expressed on Twitter about Puan Maharani. Two types of datasets were used in this study, namely an unbalanced dataset (9000 tweets: 7800 positive, 1200 negative) and a balanced dataset (2400 tweets: 1200 positive, 1200 negative). Naive Bayes classifier method is used for sentiment classification, including data collection, labeling, preprocessing, TF-IDF weighting, feature selection, data division, Naive Bayes classification, and evaluation with confusion matrix. The data is divided into 70:30, 80:20 and 90:10 ratios for training and test data. Feature selection uses a threshold of 0.001. Based on the results of the research conducted, it can be inferred that sentiment analysis can be an effective tool to understand the opinions of the public, especially netizens on the Twitter platform related to the perception of Puan Maharani. The highest accuracy value of the unbalanced dataset was obtained, which was 88.89% at a training and test data division ratio of 90:10 with the highest accuracy of the balanced dataset was 81.0% at a data division ratio of 90:10.*

Keywords: Puan Maharani; Sentiment Classification; Naive Bayes Classifier; Twitter

PENDAHULUAN

Puan Maharani ialah politisi yang sangat berpengaruh di Indonesia, khususnya di partainya, PDI Perjuangan. Selama karir politiknya, Puan telah menduduki sejumlah jabatan penting, seperti Ketua Dewan Perwakilan Rakyat (DPR) dan ketua Fraksi PDI Perjuangan. Selama masa jabatannya, ia sering diperhatikan oleh media dan masyarakat, terutama terkait dengan keputusan yang dibuatnya. Pemilihan Puan Maharani sebagai subjek penelitian ini dilatarbelakangi oleh peristiwa menarik saat anggota DPR merayakan ulang tahunnya di tengah demo kenaikan harga BBM, yang menjadi topik hangat dan menarik perhatian masyarakat, sebagaimana dilaporkan oleh CNN Indonesia."

Twitter adalah salah satu jejaring sosial terpopuler di Indonesia, dengan 18,45 juta pengguna aktif pada 2022 menggunakan platform tersebut untuk beragam informasi, berdiskusi, serta menyampaikan

pendapat mereka mengenai topik lain, termasuk politik. Twitter memiliki tingkat pembaruan yang tinggi. Hal ini menyebabkan ketersediaan informasi yang tinggi di Twitter, menjadikan Twitter tempat yang baik untuk analisis sentimen. Gudang data Twitter sangat efektif untuk penelitian di bidang politik, pemasaran, dan sosial [1].

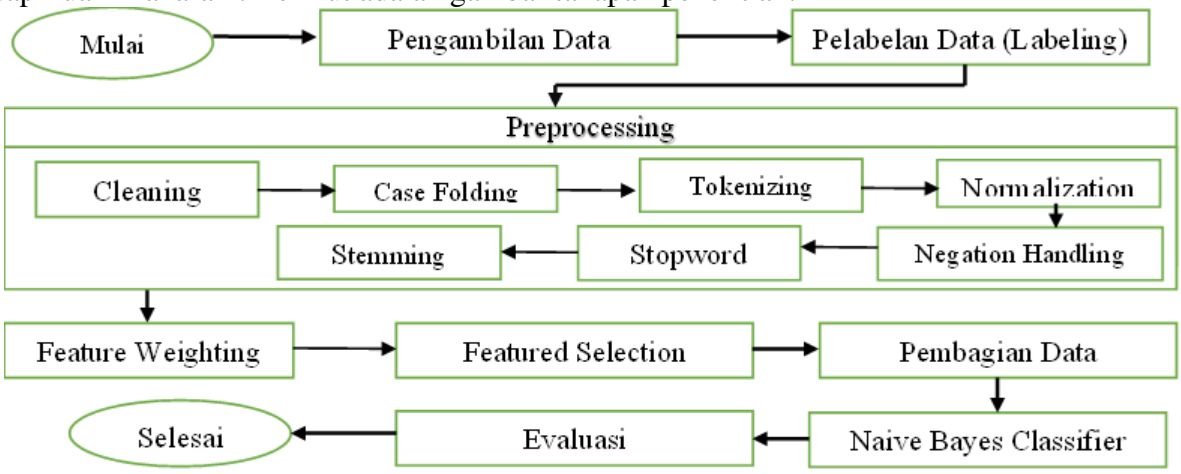
Analisis sentimen ialah teknik mengekstrak data pendapat, memahami, serta secara otomatis memproses data teks guna melihat perasaan dalam pendapat [2]. Analisis sentimen dapat menjadi alat yang efektif untuk memahami opini dan pendapat publik tentang Puan Maharani di Twitter. Analisis sentimen dapat memberikan gambaran apakah *tweet* terkait Puan Maharani memiliki sentimen positif ataupun negatif. Hal ini dapat memberikan lebih banyak wawasan tentang bagaimana masyarakat bereaksi terhadap tindakan dan keputusan politik Puan Maharani.

Terdapat dua jenis pendekatan pada klasifikasi yakni *Supervised Learning* dan *Unsupervised Learning*. Penelitian ini menggunakan *Supervised Learning*, yang menggunakan algoritma yang menciptakan fungsi yang memetakan *input* ke *output* yang dikehendaki. Pendekatan ini memiliki kemampuan untuk mengadaptasi dan mempraktikkan data pemodelan untuk tujuan serta konteks tertentu, namun membutuhkan data berlabel yang berpotensi mahal. [3].

Pada penelitian ini akan menerapkankan satu diantara sejumlah metode analisis sentimen masyarakat di Twitter yaitu Naïve Bayes Classifier. Metode Naive Bayes Classifier ialah metode klasifikasi yang cukup sederhana namun cukup akurat untuk memprediksi kelas (dalam hal ini sentimen) suatu dokumen berdasarkan kemunculan kata tertentu dalam suatu dokumen [3]. Terbukti dalam sudi yang dilaksanakan [4] yang membandingakan metode Naïve Bayes Classifier dengan Support Vector Machine (SVM) dalam mengklasifikasikan opini masyarakat di *Twitter* perihal kondisi *New Normal* di Indonesia, didapatkan nilai akurasi dari Naïve Bayes Classifier jumlah 94,55% dan akurasi dari Support Vector Machine sebesar 76,50% pada rasio perbadingan 70:30 dan juga pada penelitian analisis sentimen yang dilakukan oleh [5] perihal opini terhadap PT PAL Indonesia di Twitter dengan metode K-NN, Naïve Bayes, Decision Tree dan didapatkan nilai akurasi dari metode Naïve Bayes melebihi metode lain sebesar 84,04%. Dengan menggunakan teknik klasifikasi sentimen Naive Bayes Classifier, dapat dilakukan analisis sentimen masyarakat di Twitter terkait Puan Maharani, baik terkait dengan kebijakan politik yang dilakukannya maupun terkait isu-isu dirinya sebagai seorang politikus.

METODOLOGI PENELITIAN

Dalam penelitian ini, melibatkan 9000 data dengan menerapkan metode Naive Bayes Classifier, yang mencakup proses pengumpulan data, pelabelan data, *preprocessing*, pembobotan TF-IDF, seleksi fitur, pembagian data, klasifikasi Naive Bayes, serta evaluasi dengan *confusion matrix*. Tujuan dari penelitian ini ialah memberikan pemahaman yang lebih komprehensif mengenai persepsi masyarakat terhadap Puan Maharani serta membantu dalam menganalisis sentimen yang berkembang terhadap Puan Maharani. Berikut adalah gambar tahapan penelitian:



Gambar 1. Diagram Alur Penelitian

Pada gambar 1, terdapat alur atau tahap penelitian yang hendak dilaksanakan pada penelitian ini yakni klasifikasi sentiment dengan menerapkan metode *Naïve Bayes Classifier.*

2.1. Pengumpulan Data

Dalam rangka analisis ini, digunakan dataset yang terdiri dari 9000 data *tweet* yang diperoleh dari platform Twitter. Pengumpulan data dilaksanakan dengan memakai bahasa pemrograman Python dan eksekusi dilakukan melalui Google Colab. Data diambil mulai tanggal 30 April 2023 hingga 30 Oktober 2023.

2.2. Pelabelan Data

Proses pelabelan dilaksanakan secara manual dengan membaca *tweet* satu persatu lalu diberikan label positif atau negatif yang digunakan untuk melatih dan mengklasifikasikan data.

2.3. Preprocessing

Preprocessing adalah serangkaian langkah yang dilakukan pada teks mentah (dokumen, kalimat, atau kata-kata) untuk membersihkan, memformat, dan mengubah teks menjadi bentuk yang lebih terstruktur. Dengan *Text Preprocessing*, dapat menghilangkan *noise* dan meningkatkan kualitas data teks, memungkinkan model dan algoritma bekerja lebih baik serta menghasilkan output yang lebih akurat [6].

Berikut adalah beberapa langkah umum dalam *Text Preprocessing*:

2.3.1. Cleaning

Cleaning berfungsi untuk menghapus karakter khusus, tanda baca, serta simbol yang tidak perlu atau mengganggu dari teks, seperti tanda baca, emoji, URL, tag HTML, dan lainnya. Data cleaning bermanfaat untuk menghilangkan *noise* serta data yang tidak konsisten [7].

2.3.2. Case Folding

Case Folding ialah langkah preprocessing teks yang bertujuan untuk menjadikan semua huruf menjadi huruf kecil dalam suatu dokumen. Tujuan dari proses ini adalah untuk memudahkan pencarian dan pemrosesan teks serta mengurangi kompleksitas analisis [8].

2.3.3. Tokenizing

Tokenization ialah proses memisahkan data teks menjadi sejumlah bagian atau token [9]. Token dapat berupa kata, frasa, atau karakter tergantung pada tujuan analisis.

2.3.4. Normalization

Normalisasi ialah memperbaiki kata-kata yang yang salah pada teks merujuk korpus yang dibuat [10]. Normalisasi dilakukan untuk mengubah variasi kata menjadi bentuk standar.

2.3.5. Negation Handling

Negasi kata sendiri merupakan salah satu bentuk negasi atau negasi terhadap suatu pernyataan kata tertentu. Jika kata “antusias” digolongkan ke dalam kelas positif, maka kata-kata yang mengandung kata negatif seperti “tidak antusias” akan digolongkan ke dalam kelas negatif. Oleh karena itu, penting agar proses analisis sentimen dapat menangani negasi untuk menghindari kesalahan bias [11].

2.3.6. Stopword Removal

Stopword Removal adalah salah satu metode dapat diterapkan pada langkah prapemrosesan teks, di mana kata-kata itu akan dihapus dianggap sering muncul tetapi tidak bermakna dan mempunyai pengaruh yang signifikan terhadap makna suatu kalimat atau teks seperti "dan", "atau", "dalam", dan sebagainya [12].

2.3.7. Stemming

Pada tahap *stemming*, dilakukan perubahan kata sesuai dengan aturan yang berlaku dengan cara menghapus awalan, akhiran, dan sisipan pada dokumen atau mengubah kata-kata tersebut menjadi kata dasar [13].

e-ISSN : 2722-4368

2.4.Pembobotan TF-IDF

TF-IDF ialah kalkulasi untuk mengukur seberapa penting kata (*term*) pada dokumen beserta korpus [15]. TF-IDF terdiri dari dua komponen yaitu:

2.5.1. Term Frequency (TF)

Term Frequency adalah seberapa sering sebuah kata muncul pada sebuah dokumen. Semakin besar frekuensi kemunculannya, semakin besar bobotnya.

2.5.2. Inverse Document Frequency (IDF)

Inverse Document Frequency mengukur pentingnya sebuah kata di seluruh teks dokumen. Kata-kata yang muncul pada dokumen yang lebih sedikit cenderung lebih berbobot karena dianggap lebih informatif. *Inverse Document Frequency* dapat dihitung dengan menggunakan rumus berikut:

$$IDF = \log \frac{D}{F}$$

Keterangan:
D = jumlah dokumen dalam korpus
F = jumlah dokumen yang mengandung *term*

Setelah menghitung *term frequency* dan *inverse document frequency*, bobot TF-IDF kata-kata tertentu dalam dokumen diperoleh dengan mengalikan nilai TF dengan nilai IDF:

$$TF - IDF = TF \times IDF$$

Keterangan:
TF = Term Frequency
IDF = Inverse Document Frequency

2.5. Feature Selection

Feature selection dalam klasifikasi sentimen merujuk pada proses pemilihan fitur atau atribut yang paling relevan dari dataset untuk digunakan dalam membangun model klasifikasi sentimen. Tujuannya adalah untuk mengurangi atribut yang kurang relevan dan meningkatkan akurasi serta kinerja model klasifikasi [16].

2.6.Pembagian Data

Data dalam klasifikasi sentiment dibagi menjadi dua bagian: data pelatihan serta data tes. Data pelatihan diterapkan guna mengembangkan model klasifikasi, sementara data tes diterapkan untuk menilai kinerja model. Data pelatihan dan tes dapat dibagi menjadi berbagai rasio, seperti 80:20 atau 70:30 [14].

2.7. Naive Bayes Classifier

Naive Bayes classifier ialah pendekatan klasifikasi probabilitas dasar dengan ketergantungan (independen) tinggi yang menggunakan teorema Bayes [17]. Naive Bayes Classifier bekerja dengan menghitung probabilitas setiap fitur dalam data untuk setiap kemungkinan kelas. Kemudian, dengan menggunakan teorema Bayes, algoritma ini menghitung probabilitas kelas untuk data berdasarkan probabilitas fitur dalam data. Berikut tahapan algoritma Naïve Bayes Classifier:

2.7.1. Menghitung probabilitas bersyarat/likelihood:

$$P\ (x|C) = P(x1, x2, \dots, xn|C)$$

Keterangan :
C = Class
P(xi|C) = proporsi dokumen dari class C yang mengandung nilai atribut xi

2.7.2. Menghitung probabilitas prior untuk setiap kelas

$$P(C) = \frac{Nj}{N}$$

Keterangan:
Nj = jumlah dokumen pada suatu class
N = jumlah total dokumen

2.7.3. Menghitung probabilitas posterior dengan rumus:

$$P(C|x) = \frac{p(x|C)p(c)}{p(x)}$$

2.8. Evaluasi

Evaluasi model digunakan dalam klasifikasi sentimen untuk menentukan jumlah akurasi dalam data klasifikasinya. Untuk mengukur keakuratan atau ketidaktepatan data yang diklasifikasikan, teknik evaluasi model biasanya menggunakan *confusion matrix* yang dapat menghasilkan akurasi, *recall*, *precision*, serta *f1 score* [18]. Pengujian *confusion matrix* ditunjukkan pada tabel berikut:

Table 1. Confusion Matrix

Dari tabel 1 bisa didapatkan nilai *Accuracy*, *reccal, precision* dan *f1 score* dengan rumus sebagai berikut:

a) *Accuracy* (Akurasi): Rasio dari total prediksi yang benar dengan total data.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

	Actual Positive	Actual Negative
Predicted Positive	True Positive	False Negative
Predicted Negative	False Positive	True Negative

b) *Precision* (Presisi): Rasio dari kasus positif yang terprediksi benar dengan total prediksi positif.

$$Precision = \frac{TP}{TP + FP}$$

c) *Recall* (Sensitivity): Rasio dari kasus positif yang terprediksi benar dengan total kasus positif aktual.

$$Recall = \frac{TP}{TP + FN}$$

d) *F1 Score*: Rata-rata harmonik antara precision dan recall.

$$F1\ Score = \frac{2 \times Recall \times Precision}{Recall + Precision}$$

HASIL DAN PEMBAHASAN

3.1.Pengumpulan Data

Proses pengumpulan data *tweet* Twitter tentang Puan Maharani dilakukan dengan memakai Bahasa pemrograman *python* yang dijalankan di google colab yang. Berikut contoh dari *tweet* yang didapatkan.

Table 2. Contoh Tweet

Tweet	username
@mariookam395 @puanmaharani_ri Pastinya Ibu Puan layak jadi presiden RI	BellaB6225355
@puanmaharani_ri Selalu bekerja untuk memudahkan dan menguntungkan rakyat. Terimakasih Bu Puan!	AdrianSitorus95

3.2.Pelabelan Data

Pada fase ini, data *tweet* yang didapatkan akan diberikan label positif atau negatif secara manual. Proses Pelabelan dilakukan seorang ahli Bahasa Indonesia yaitu Gadis Sari Elin S.Pd. Setelah dilakukan pelabelan didapatkan 7800 data *tweet* positif dan 1200 data *tweet* negatif.

3.2. Text Preprocessing

Berikut adalah contoh hasil dari tahapan *text preprocessing*.

Table 3. Contoh Tahapan Text Preprocessing

Tahapan	Sebelum	Sesudah
Cleaning	@puan_maharani Kedekatan hubungan Indonesia-Aljazair sdh tdk diragukan lagi sejak zaman Bung Karno dan kini diteruskan oleh Ibu Puan	Kedekatan hubungan Indonesia Aljazair sdh tdk diragukan lagi sejak zaman Bung Karno dan kini diteruskan oleh Ibu Puan
Case Folding	Kedekatan hubungan Indonesia Aljazair sdh tdk diragukan lagi sejak zaman Bung Karno dan kini diteruskan oleh Ibu Puan	kedekatan hubungan indonesia aljazair sdh tdk diragukan lagi sejak zaman bung karno dan kini diteruskan oleh ibu puan
Tokenizing	kedekatan hubungan indonesia aljazair sdh tdk diragukan lagi sejak zaman bung karno dan kini diteruskan oleh ibu puan	[kedekatan,hubungan,indonesia,aljazair, sdh,tdk,diragukan,lagi,sejak,zaman, bung,karno,dan,kini,diteruskan,oleh,ibu ,puan]
Normalization	[kedekatan,hubungan,indonesia, aljazair,**sdh**,**tdk**,diragukan,lagi, sejak,zaman,bung,karno,dan,kini, diteruskan,oleh,ibu ,puan]	[kedekatan,hubungan,indonesia,aljazair, **sudah**,**tidak**,diragukan,lagi,sejak,zaman, bung,karno,dan,kini,diteruskan,oleh,ibu ,puan]
Negation Handling	[kedekatan,hubungan,indonesia, aljazair,sudah,**tidak**,**diragukan**,lagi ,sejak,zaman,bung,karno,dan,kini, diteruskan,oleh,ibu ,puan]	[kedekatan,hubungan,indonesia,aljazair, sudah,**percaya**,lagi,sejak,zaman,bung, karno,dan,kini,diteruskan,oleh,ibu ,puan]
Stopword Removal	[kedekatan,hubungan,indonesia, aljazair,sudah,**percaya**,lagi, sejak,zaman,bung,karno,dan,kini, diteruskan,oleh,ibu ,puan]	[kedekatan,hubungan,indonesia,aljazair, percaya,lagi,sejak,zaman,bung,karno,kini ,diteruskan,ibu,puan]
Stemming	**kedekatan**,**hubungan**,indonesia,alj azair,percaya,lagi,sejak,zaman,bun g,karno,kini,**diteruskan**,ibu,puan]	**dekat**,**hubung**,indonesia,aljazair,percaya ,lagi,sejak,zaman,bung,karno,kini,**terus**, ibu,puan]

3.3.Pembobotan TF-IDF

Pada tahap TF-IDF, digunakan library scikit-learn(sklearn) dengan modul TfidfVectorizer.

3.4.Feature Selection

Setelah dilaksanakan pembobotan dengan TF-IDF, dilakukan pemilihan fitur. Pada penelitian ini pemilihan fitur menggunakan *threshold* 0,001.

3.5.Pembagian Data

Pada fase pembagian data, data *tweet* yang ada akan dibagi menjadi data latih dan data uji berdasarkan rasio yakni 70:30, 80:20 dan 90:10.

3.6.Naïve Bayes Classifier

Setelah dilakukan *feature selection*, dilaksanakan proses klasifikasi dengan menerapkan metode Naive Bayes Classifier dengan rasio data latih serta data uji 70:30, 80:20, dan 90:10. Terdapat 2 jenis dataset yang digunakan dalam penelitian ini ialah data tidak seimbang (*Unbalance*) serta data seimbang (*balance*). Data tidak seimbang yang dimaksud ialah data yang jumlah *tweet* yang berlabel positif dan negatif tidak seimbang sedangkan data seimbang sebaliknya. Terdapat 9000 *tweet* pada data tidak seimbang serta 2000 *tweet* untuk data yang seimbang yang sudah diberi label positif dan negatif.

3.7.Evaluasi

Selanjutnya dilakukan evaluasi dari hasil klasifikasi sentimen dari setiap rasio pembagian data dan didapatkan *confusion matrix* sebagai berikut.

Table 4. Hasil Confusion Matrix

Confusion Matrix						
	Data tidak Seimbang			Data Balance		
	70:30	80:20	90:10	70:30	80:20	90:10
TP	2347	1567	790	269	178	92
TN	27	22	10	283	199	92
FP	320	205	97	67	41	22
FN	6	6	3	77	46	26

Dari tabel *confusion matrix* diatas, didapatkan nilai *accuracy, precision, recall* serta *f1-score* untuk setiap pengujian sebagai berikut:

Table 5. Nilai Accuracy

	Data Tidak Seimbang			Data Balance		
	70:30	80:20	90:10	70:30	80:20	90:10
Precision	87,93%	88,3%	88,89%	79,0%	80,4%	81,0%
Recall	87,93%	88,3%	88,89%	79,0%	80,4%	81,0%
F1-Score	87,93%	88,3%	88,89%	79,0%	80,4%	81,0%
Accuracy	87,93%	88,3%	88,89%	79,0%	80,4%	81,0%

Dari tabel 5, nilai akurasi yang tertinggi dari data *unbalance* sebesar 88,89% pada rasio pembagian data 90:10 dan nilai akurasi tertinggi dari data *balance* sebesar 81,0 pada rasio pembagian data 90:10. Akurasi dan metrik lainnya cenderung lebih tinggi pada data yang tidak seimbang (sekitar 87–88%) dibandingkan dengan data yang seimbang (sekitar 79–81%). Ini menunjukkan bahwa model mungkin memberikan hasil yang lebih baik pada data yang tidak seimbang, mungkin karena dominasi kelas mayoritas yang mempengaruhi hasil klasifikasi. Namun, meskipun akurasi lebih tinggi, menggunakan data yang seimbang tetap penting untuk memastikan bahwa model tidak bias.

3.1.Wordcloud

Wordcloud adalah alat visualisasi data teks yang biasa digunakan untuk menampilkan kata-kata yang paling sering muncul dalam teks [8]. *Wordcloud* menunjukkan frekuensi dan warna kata yang bisa digunakan untuk menunjukkan sentimen positif, negatif, atau netral. Berikut hasil *wordcloud* dari sentimen terhadap Puan Maharani setelah preprocessing

e-ISSN : 2722-4368

Gambar 2. Wordcloud data tidak seimbang

Gambar 3. Wordcloud data seimbang

KESIMPULAN

Berdasarkan hasil penelitian yang dilaksanakan, dapat disimpulkan bahwa analisis sentimen dapat menjadi alat yang efektif untuk memahami pendapat masyarakat khususnya masyarakat di platform Twitter terkait dengan persepsi terhadap Puan Maharani. Nilai akurasi tertinggi dari proses pengujian data *unbalance* didapatkan yaitu sebesar 88.89% pada rasio pembagian data latih serta data uji 90:10 dan nilai akurasi dari pengujian data *balance* sebesar 81,0% pada rasio pembagian data 90:10. Namun ditemukan bahwa ketika menguji data yang tidak seimbang, pengklasifikasi Naive Bayes cenderung memihak pada kelas mayoritas (positif), sehingga dapat menyebabkan kesalahan dalam mengklasifikasikan *tweet* negatif. Untuk penelitian selanjutnya, teknik untuk mengatasi bias ini harus dipertimbangkan, seperti menggunakan metode keseimbangan kelas atau model yang lebih kompleks.

DAFTAR PUSTAKA

[1] F. Yuspriyadi, "Klasifikasi Sentimen Twitter Menggunakan Lstm," *METHODIKA: Jurnal Teknik Informatika dan Sistem Informasi*, vol. 9, no. 1, pp. 4–8, 2023, doi: 10.46880/mtk.v9i1.1720.

[2] F. V. Sari and A. Wibowo, "Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi," *Jurnal SIMETRIS*, vol. 10, no. 2, pp. 681–686, 2019.

[3] T. S. Az-Zahra, "Analisis sentimen terhadap belajar daring menggunakan optimasi naive bayes classifier dengan adaboost," UIN Sunan Ampel Surabaya, 2021. [Online]. Available: digilib.uinsby.ac.id

[4] E. A. Lisangan, A. Gormantara, and R. Y. Carolus, "Implementasi Naive Bayes pada Analisis Sentimen Opini Masyarakat di Twitter Terhadap Kondisi New Normal di Indonesia," *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, vol. 2, no. 1, pp. 23–32, 2022, doi: 10.24002/konstelasi.v2i1.5609.

[5] F. S. Pattiiha and H. Hendry, "Perbandingan Metode K-NN, Naïve Bayes, Decision Tree untuk Analisis Sentimen Tweet Twitter Terkait Opini Terhadap PT PAL Indonesia," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, no. 2, p. 506, 2022, doi: 10.30865/jurikom.v9i2.4016.

[6] M. Yasid, "Analisis Sentimen Maskapai Citilink Pada Twitter Dengan Metode Naïve Bayes," *JURNAL ILMIAH INFORMATIKA*, vol. 7, no. 02, p. 82, Oct. 2019, doi: 10.33884/jif.v7i02.1329.

[7] A. Herdhianto, "Sentiment Analysis Menggunakan Naïve Bayes Classifier (NBC) Pada Tweet Tentang Zakat," Universitas Islam Negeri Syarif Hidayatullah, 2020.

[8] A. Irsyad and R. D. Geralda, "Analisis Sentimen SEA Games 2023 di Twitter Metode dengan Machine Learning," *Adopsi Teknologi dan Sistem Informasi (ATASI)*, vol. 2, no. 2, pp. 126–131, 2023, doi: 10.30872/atasi.v2i2.1138.

[9] D. Aditya, A. Mubarok, M. Kom, and S. Susanti, "Analisis Sentimen Pada Media Sosial Twitter Menggunakan Metode Naïve Bayes Classifier (Studi Kasus: Komentar Publik Kepada Tri Indonesia)," *Jurnal Informatika*, vol. 6, no. 1, pp. 1–8, 2019, [Online]. Available: http://ejournal.bsi.ac.id/ejurnal/index.php/ji/article/view/-

[10] S. N. J. Fitriyyah, N. Safriadi, and E. E. Pratama, "Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 5, no. 3, p. 279, 2019, doi: 10.26418/jp.v5i3.34368.

[11] R. I. Tarecha, F. Wahyudi, and U. M. Jannah, "Penanganan Negasi dalam Analisa Sentimen Bahasa Indonesia," *Jurnal Sistem Informasi dan Informatika (JUSIFOR)*, vol. 1, no. 1, pp. 51–58, 2022, doi: 10.33379/jusifor.v1i1.1276.

[12] S. J. Angelina, A. Bijaksana, P. Negara, and H. Muhardi, "Analisis Pengaruh Penerapan Stopword Removal Pada Performa Klasifikasi Sentimen Tweet Bahasa Indonesia," *JUARA (Jurnal Aplikasi dan Riset Informatika)*, vol. 02, no. 1, pp. 165–173, 2023, doi: 10.26418/juara.v2i1.69680.

[13] M. M. Mala Olhang, S. Achmadi, and F. X. A. Wibisono, "ANALISIS SENTIMEN PENGGUNA TWITTER TERHADAP COVID-19 DI INDONESIA MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER (NBC)," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 4, no. 2, pp. 214–221, Dec. 2020, doi: 10.36040/jati.v4i2.2695.

[14] A. Nofandi, N. Y. Setiawan, and D. W. Brata, "Analisis sentimen ulasan pelanggan dengan Metode Support Vector Machine (SVM) untuk peningkatan kualitas layanan pada Restoran Warung Wareg," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 1, pp. 458–466, 2023, [Online]. Available: https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/12218

[15] E. Febriyani and H. Februariyanti, "Analisis sentimen terhadap program kampus merdeka menggunakan algoritma naive bayes classifier di twitter," *Jurnal Tekno Kompak*, vol. 17, no. 1, pp. 25–38, 2023, [Online]. Available: https://ejurnal.teknokrat.ac.id/index.php/teknokompak/article/view/2061

[16] F. Septianingrum and A. S. Y. Irawan, "Metode Seleksi Fitur Untuk Klasifikasi Sentimen Menggunakan Algoritma Naive Bayes: Sebuah Literature Review," *Jurnal Media Informatika Budidarma*, vol. 5, no. 3, p. 799, 2021, doi: 10.30865/mib.v5i3.2983.

[17] L. A. Andika, P. A. N. Azizah, and R. Respatiwulan, "Analisis Sentimen Masyarakat terhadap Hasil Quick Count Pemilihan Presiden Indonesia 2019 pada Media Sosial Twitter Menggunakan Metode Naive Bayes Classifier," *Indonesian Journal of Applied Statistics*, vol. 2, no. 1, p. 34, 2019, doi: 10.13057/ijas.v2i1.29998.

[18] R. L. Atimi and Enda Esyudha Pratama, "Implementasi Model Klasifikasi Sentimen Pada Review Produk Lazada Indonesia," *Jurnal Sains dan Informatika*, vol. 8, no. 1, pp. 88–96, 2022, doi: 10.34128/jsi.v8i1.419.