

PENERAPAN METODE *MULTINOMIAL NAÏVE BAYES* UNTUK MENGANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI WAHYOO

Ilham Alamsyah^{1*}, Rizky Tahara Shita²

^{1,2}Teknik Informatika, Fakultas Teknologi Informasi, Universitas Budi Luhur, Jakarta, Indonesia

Email: ¹ilhamalamsyah510@gmail.com, ²rizky.taharashita@budiluhur.ac.id

(* : *corresponding author*)

Abstrak-Aplikasi Wahyoo adalah penyedia jaringan distribusi dan penjualan yang bertujuan untuk meningkatkan efisiensi dan potensi UMKM di bidang kuliner yang saat ini dalam jangkauan wilayah Jabodetabek. Wahyoo menyediakan lebih dari 2.000 jenis bahan baku, termasuk beras, minyak goreng, tepung, dan produk segar seperti sayuran, buah, dan daging. Selain itu, Wahyoo juga menyediakan produk siap jual seperti FMCG (Fast Moving Consumer Goods) dan produk siap masak seperti Frozen Food. Dalam pelaksanaannya, aplikasi wahyoo membantu pengguna dan memiliki keuntungan lain yang membuat pengguna merasa puas terhadap layanan yang diberikan. Namun, aplikasi ini juga memiliki masalah atau fitur yang kurang disukai oleh penggunanya. Terkait tinjauan pengguna mengenai kepuasan atau ketidakpuasan penggunaan aplikasi wahyoo, salah satunya adalah dengan memberikan tinjauan di Google Play Store. Penelitian ini bertujuan untuk melakukan analisis sentimen pengguna terhadap layanan aplikasi dengan pendekatan machine learning dan ekstraksi fitur CountVectorizer dan Algoritme Multinomial Naive Bayes. Dataset yang digunakan bersumber dari ulasan pengguna aplikasi pada Google Play Store, hasil analisis berdasarkan 413 ulasan yang diterima menunjukkan nilai positif ulasan memiliki angka yang lebih tinggi dibandingkan dengan nilai negatif ulasan dengan persentase 71,7%. Hasil pengujian dari dataset tersebut didapatkan nilai akurasi sebesar 88%, presisi sebesar 90%, dan recall sebesar 92%.

Kata Kunci: Ulasan, Klasifikasi, *Multinomial Naïve Bayes*, Ekstraksi Fitur

IMPLEMENTATION OF THE MULTINOMIAL NAÏVE BAYES METHOD TO ANALYZE SENTIMENT OF USER REVIEWS OF WAHYOO APPLICATION

Abstract-The Wahyoo application is a distribution and sales network provider that aims to increase the efficiency and potential of MSMEs in the culinary field which are currently within the reach of the Jabodetabek area. Wahyoo provides more than 2,000 types of raw materials, including rice, cooking oil, flour and fresh products such as vegetables, fruit and meat. In addition, Wahyoo also provides ready-to-sell products such as FMCG (Fast Moving Consumer Goods) and ready-to-cook products such as Frozen Food. In its implementation, the Wahyoo application helps its users and has other advantages that make users feel satisfied regarding the services provided. However, this application is also not free from problems or errors faced by its users. Regarding user reviews on satisfaction or dissatisfaction with the use of the Wahyoo application, one of which is by providing reviews on the Google Play Store. This research aims to examine user emotions towards application services using a machine learning technique and feature extraction CountVectorizer and Algoritme Multinomial Naive Bayes. The dataset used comes from application user reviews on the Google Play Store, the findings of the examination, relying on 413 received evaluations, indicate that the sentiment of positive reviews holds more significance than that of negative reviews, with a proportion of 71.7%. The outcomes of the experiment, derived from the set of data, yielded an accuracy rate of 88%, a precision rate of 90%, and a recall rate of 92%.

Keywords: Reviews, Classification, *Multinomial Naïve Bayes*, Feature Extraction

1. PENDAHULUAN

Aplikasi Wahyoo adalah penyedia jaringan distribusi dan penjualan yang bertujuan untuk meningkatkan efisiensi dan potensi UMKM di bidang kuliner yang saat ini dalam jangkauan wilayah Jabodetabek. Dalam pelaksanaannya, aplikasi wahyoo membantu pengguna dan memiliki keuntungan tambahan yang membuat pengguna merasa puas terkait layanan yang diberikan. Namun, aplikasi ini juga tidak terhindar dari permasalahan atau kesalahan yang dihadapi oleh penggunanya, hal tersebut dapat menimbulkan ketidakpuasan pengguna.

Berkaitan dengan ulasan pengguna mengenai kepuasan atau ketidakpuasan penggunaan aplikasi wahyoo, salah satu caranya adalah dengan memberikan ulasan di *Google Play Store*. Dari ulasan tersebut, diperlukan proses untuk membedakan pendapat yang positif dan negatif dengan melakukan analisis sentimen. Dari banyaknya *review* yang tersedia, jika dilakukan analisis akan berguna bagi Wahyoo dalam mengkategorikan sentimen positif atau

negatif serta hasil sentimen bisa digunakan sebagai informasi dalam mengevaluasi perkembangan aplikasi tersebut.

Analisis sentimen adalah teknik untuk mengekstrak informasi opini, memahami, dan memproses teks secara otomatis untuk mengidentifikasi emosi yang terkandung dalam opini. Pengguna internet melakukan analisis pendapat di platform media sosial untuk memberikan penilaian atau pandangan pribadi. Tujuan dari analisis sentimen adalah untuk menentukan pendapat seseorang dengan mempertimbangkan topik yang spesifik [1]. Pada penelitian sebelumnya mengenai analisis sentimen terhadap kandidat presiden Indonesia 2019 dari media sosial twitter yang menggunakan metode naïve bayes [2]. Penelitian tersebut menyimpulkan bahwa dalam pengujian 3 kategori kandidat presiden 01 dan 02, diperoleh tingkat ketepatan masing-masing sebesar 64,6% dan 58%. Sementara itu, pada pengujian 2 kategori kandidat presiden 01 dan 02, ditemukan hasil ketepatan berturut-turut sebesar 77,7% dan 88%. Kandidat presiden nomor urut dua menunjukkan kinerja tertinggi dimana nilai f-measure mencapai 0,88.

Metode naïve bayes dapat juga diimplementasikan dalam menganalisis sentimen opini publik mengenai larangan mudik pada twitter dan didapatkan nilai akurasi yaitu 83,38% [3]. Dengan latar belakang tersebut, tujuan dari penelitian ini adalah untuk melakukan evaluasi sentimen dari ulasan pengguna kepada aplikasi Wahyoo. Dilakukan penelitian sentimen pada ulasan aplikasi Wahyoo dengan metode klasifikasi multinomial naïve bayes.

Metode multinomial naïve bayes dipilih sebab diketahui mempunyai akurasi tinggi meskipun perhitungannya sederhana [4]. Selain itu, metode multinomial naïve bayes menerapkan dasar-dasar probabilitas kata per kata dan label kategori untuk memprediksi probabilitas label pada dokumen, umumnya digunakan pada penelitian pengklasifikasian teks karena efisiensi dan kemudahan perhitungannya [4]. Tujuan dari penelitian ini adalah untuk memahami alur kerja metode multinomial naïve bayes pada analisis sentimen pada ulasan aplikasi Wahyoo di Google Play Store dan mendapatkan pemahaman yang lebih baik tentang kepuasan pengguna yang digunakan untuk fitur yang dikembangkan.

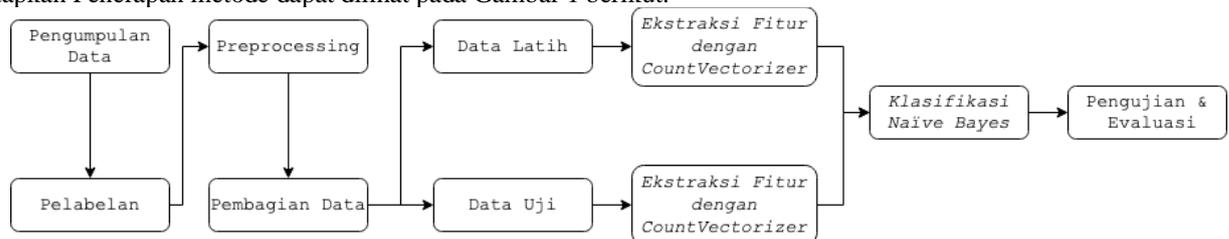
2. METODE PENELITIAN

2.1 Data Penelitian

Penelitian ini menggunakan dataset yang berasal dari ulasan pengguna aplikasi wahyoo pada *google play store*. Dataset diperoleh melalui fase pengumpulan data dibantu menggunakan *package google-play-scraper* dengan mensortir berdasarkan ulasan yang paling relevan.

2.2 Penerapan Metode

Dalam membuat sistem analisis sentimen untuk data review aplikasi Wahyoo di Google Play Store dengan menggunakan metode multinomial naïve bayes, dibutuhkan gambaran rencana implementasi yang digunakan dalam penelitian ini. Gambaran ini dimaksudkan agar penelitian bisa dilaksanakan sesuai rencana yang telah ditetapkan Penerapan metode dapat dilihat pada Gambar 1 berikut:



Gambar 1. Tahapan Metode

2.3 Pengumpulan Data

Pengumpulan dilakukan data dibantu dengan *package google-play-scraper* dengan mensortir berdasarkan ulasan yang paling relevan. Selanjutnya mengurutkan data berdasarkan tanggal dibuatnya ulasan secara descending. Setelah proses pemilahan dan pengurutan tanggal dibuatnya selesai, maka berikutnya yaitu mendownload dataset tersebut ke dalam format csv.

2.4 Pelabelan

Data yang sudah terhimpun selanjutnya diberi label sentimen untuk setiap data tinjauan. Semua tinjauan yang terkumpul selanjutnya diberi label positif atau label negatif yang dilaksanakan secara manual. Label positif

diberikan untuk tinjauan yang membangun, ungkapan terima kasih, dan penghargaan. Label negatif diberikan untuk tinjauan seperti keluhan, ejekan, dan mencemarkan nama baik kepada aplikasi.

2.5 Preprocessing

Data yang diberikan label selanjutnya dilakukan *preprocessing*. Preprocessing dilakukan pada data teks yang akan diubah menjadi format yang lebih baik sehingga menghasilkan informasi teks dengan mutu yang prima dan siap dimanfaatkan pada proses berikutnya [5]. Dalam penelitian ini ada beberapa tahapan pada preprocessing sebagai berikut:

a. *Case Folding*

Case Folding adalah penyamaan *case* dalam dokumen [6]. *Case folding* bertujuan untuk menyamaratakan seluruh kata.

b. *Cleansing*

Cleansing adalah proses menghapus karakter dan simbol yang tidak diperlukan pada suatu teks seperti tanda pagar (#), mentions (@), bilangan, tanda baca, emoji, dan menghilangkan spasi yang berlebihan.

c. Mengubah *Slang Word*

Mengubah slang word merupakan proses menggantikan kata-kata informal menjadi kata-kata formal. Tindakan ini diterapkan supaya kata-kata yang disingkat atau diperpanjang dapat diubah menjadi kata-kata standar yang sesuai dengan Kamus Besar Bahasa Indonesia (KBBI) [7].

d. *Tokenization*

Tokenization adalah memecah kata-kata dalam kalimat, paragraf, atau halaman menjadi potongan kata tunggal atau *termmed word* yang berdiri sendiri [8].

e. *Stopword Removal*

Stopword removal merupakan proses mengekstraksi kata-kata penting dari hasil pemisahan kata. Terdapat algoritma stoplist yang berusaha menghapus kata-kata yang tidak diperlukan dan daftar kata yang berfungsi sebagai tempat penyimpanan kata-kata penting [9]. Kata-kata yang dihapus adalah kata-kata yang tidak deskriptif, seperti yang, dan, di dan pada.

f. *Stemming*

Stemming merupakan sebuah proses di mana kata-kata yang telah melewati proses filtering diubah menjadi bentuk dasarnya dengan menghilangkan imbuhan pada dokumen [9]. Sebagai contoh, "mempertahankan" diubah menjadi "tahan", "mengakui" menjadi "akui", dan "melindungi" menjadi "lindung".

2.6 Pembagian Data

Dataset yang sudah diberi label kemudian dipecah menjadi bagian-bagian data. Dataset dibagi dengan rasio 80:20 untuk setiap data latih dan data uji.

2.7 CountVectorizer

Mesin tidak dapat mengerti karakter dan kata-kata. Oleh karena itu, ketika berhadapan dengan data teks, kita harus mengubahnya menjadi angka agar alat dapat memahaminya. *CountVectorizer* adalah proses pengolahan dokumen atau teks yang mengubahnya menjadi vektor. *CountVectorizer* digunakan untuk menghitung jumlah kata dalam dokumen atau kalimat dan kemudian merepresentasikannya dalam bentuk vektor [10].

2.8 Multinomial Naïve Bayes

Multinomial Naïve Bayes merupakan teknik klasifikasi data probabilistik sederhana yang menggunakan teorema Bayes dan mengasumsikan independensi antara fitur individual. Dalam metode ini, setiap dokumen direpresentasikan dengan kata-kata yang disimbolkan sebagai " $a_1, a_2, a_3, \dots, a_n$ ", dimana a_1 merupakan kata pertama dan seterusnya hingga kata ke-n. Pada metode ini entri kelas dilambangkan dengan V . Kemudian dilakukan pencarian nilai tertinggi dari kelas teks yang diuji (V_{MAP}). Persamaan V_{MAP} dapat dilihat pada persamaan 1.

$$V_{MAP} = \underset{v_i \in V}{argmax} P(V_j) \cap_i P(a_i | v_j) \quad (1)$$

Nilai $P(v_j)$ adalah probabilitas bahwa dokumen tersebut akan diklasifikasikan dalam kategori j , terlepas dari isinya. Nilai $P(v_j)$ diperoleh dari persamaan 2.

$$P(v_j) = \frac{|Docs_j|}{|training|} \quad (2)$$

Keterangan:

|Docs j| : Banyaknya dokumen dengan kategori j pada dokumen latih
 |training| : Banyaknya dokumen data latih

Nilai $P(a_i|v_j)$ adalah peluang memiliki kata a_i dalam pada kelas v_j . Nilai $P(a_i|V_j)$ $P(a_i|v_j)$ dihasilkan dari Persamaan 3.

$$P(a_i|v_j) = \frac{|n_i + 1|}{|n + kosakata|} \quad (3)$$

Keterangan:

n_i : Frekuensi munculnya kata a_i dalam dokumen latih dengan kategori v_j
 n : Banyaknya seluruh kata dalam dokumen latih dengan kategori v_j
 v_j kosakata : Total kata unik dalam seluruh dokumen latih

2.9 Pengujian

Dalam penelitian ini pengujian dilakukan terhadap model pelatihan yang menggunakan metode Multinomial Naïve Bayes dengan tujuan untuk mengetahui tingkat keakuratan, presisi, dan recall. Pengujian dilakukan menggunakan *confusion matrix*. *Confusion Matrix* merupakan tabel yang menampilkan hasil klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah [9]. Tabel *confusion matrix* dapat dilihat pada Tabel 1 berikut:

Tabel 1. *Confusion Matrix*

	Predicted Positif	Predicted Negatif
Actual Positif	TP	FN
Actual Negatif	FP	TN

Keterangan :

TP (True Positive) : Total dokumen dengan label positif yang tepat terklasifikasi sebagai label positif.
 TN (True Negative) : Total dokumen label negatif yang tepat terklasifikasi sebagai label negatif.
 FP (False Positive) : Total dokumen label negatif yang salah terklasifikasi sebagai label positif.
 FN (False Negative) : Total dokumen label positif yang terklasifikasi salah menjadi label negatif.

2.10 Akurasi

Akurasi merupakan rasio prediksi Benar (positif dan negatif) pada keseluruhan data. Perhitungan nilai akurasi dihasilkan dari Persamaan 4.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

2.11 Presisi

Presisi merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Perhitungan nilai presisi dihasilkan dari Persamaan 5.

$$Presisi = \frac{TP}{TP + FP} \quad (5)$$

2.12 Recall

Recall merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Perhitungan nilai *recall* dihasilkan dari Persamaan 6.

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

3. HASIL DAN PEMBAHASAN

3.1 Tahap Pengumpulan Data

Pengumpulan data dibantu oleh *package google-play-scraper* dengan mensortir berdasarkan ulasan yang paling relevan dan didapatkan *dataset* sebanyak 413 data. Kemudian mengurutkan data berdasarkan tanggal dibuatnya ulasan secara descending. Setelah proses pemilahan dan pengurutan tanggal dibuatnya selesai, maka berikutnya yaitu mendownload dataset tersebut ke dalam format csv. Tabel sampel *dataset* dapat dilihat pada Tabel 2 berikut:

Tabel 2. Sampel Dataset

Content	Score
Suka bgt... apalagi jaman ada pay later.. sayang aj skrg jadi d kenai ongkir... kudu belanja banyak dulu baru bisa freeong... 😞 Case Folding	3
suka bgt... apalagi jaman ada pay later.. sayang aj skrg jadi d kenai ongkir... kudu belanja banyak dulu baru bisa freeong... 😞 Cleansing	2
suka bgt apalagi jaman ada pay later sayang aj skrg jadi d kenai ongkir kudu belanja banyak dulu baru bisa freeong Semakin kesini semakin mahal bgt..	2
Aplikasi yang sangat membantu kebutuhan memasak	5

3.2 Tahap Pelabelan

Tahap pelabelan dilakukan secara manual. Tinjauan yang bersifat membangun, ungkapan terima kasih, dan apresiasi akan diberikan label positif. Sementara itu, Tinjauan yang berupa keluhan, ejekan, dan merendahkan aplikasi akan diberikan label negatif.

3.3 Tahap Preprocessing

Preprocessing dilakukan terhadap dataset melalui beberapa tahapan yaitu *case folding*, *cleansing*, mengganti *slang word*, *tokenizing*, *stopword removal*, *stemming*. Tabel tahap *preprocessing* dapat dilihat pada Tabel 3 berikut:

Tabel 3. Tahap Preprocessing

Proses	Hasil Preprocessing
Ulasan Awal	Suka bgt... apalagi jaman ada pay later.. sayang aj skrg jadi d kenai ongkir... kudu belanja banyak dulu baru bisa freeong... 😞
Case Folding	suka bgt... apalagi jaman ada pay later.. sayang aj skrg jadi d kenai ongkir... kudu belanja banyak dulu baru bisa freeong... 😞
Cleansing	suka bgt apalagi jaman ada pay later sayang aj skrg jadi d kenai ongkir kudu belanja banyak dulu baru bisa freeong
Mengganti Slang Word	suka banget apalagi jaman ada pay later sayang saja sekarang jadi di kenai ongkos kirim harus belanja banyak dulu baru bisa freeong
Tokenization	[suka, banget, apalagi, jaman, ada, pay, later, sayang, saja, sekarang, jadi, di, kenai, ongkos, kirim, harus, belanja, banyak, dulu, baru, bisa, freeong]
Stopword Removal	[suka, banget, jaman, pay, later, sayang, kenai, ongkos, kirim, belanja, freeong]
Stemming	suka banget jaman pay later sayang kena ongkos kirim belanja freeong

3.4 Tahap Pembagian Data

Ulasan yang selesai dipreprocessing kemudian dibagi menjadi dua bagian, yaitu data latih dan data uji. Pembagian data dilaksanakan dengan menggunakan rasio yang telah ditetapkan, yaitu 80:20 atau 80% data latih dan 20% data uji. Dalam penelitian ini, proses pembagian data dilakukan pada 413 data ulasan.

3.5 Tahap Ekstraksi Fitur *CountVectorizer*

Ekstraksi fitur dilaksanakan dengan menggunakan metode *CountVectorizer* pada penelitian ini. Pada tahap ekstraksi fitur, teks uji ulasan diubah menjadi angka, yang selanjutnya dapat diproses pada tahap klasifikasi *multinomial naïve bayes*. Tabel data latih dapat dilihat pada Tabel 4 berikut:

Tabel 4. Data Latih

Content	Label
[senang, belanja, aplikasi, wahyoo]	Positif
[belanja, mudah]	Positif
[belanja, buruk, mahal]	Negatif

Tabel 4 adalah contoh data pelatihan yang digunakan dalam penelitian ini yang selanjutnya melibatkan langkah ekstraksi fitur. Tabel ekstraksi fitur dapat dilihat pada Tabel 5 berikut:

Tabel 5. Ekstraksi Fitur

Kata	Positif	Negatif
aplikasi	1	0
Belanja	2	1
Buruk	0	1
Mahal	0	1
Mudah	1	0
Senang	1	0
Wahyoo	1	0
Total kata berlabel positif	9	
Total kata berlabel negatif		4
Total kosakata	11	

3.6 Tahap Klasifikasi Multinomial Naïve Bayes

Tahapan klasifikasi multinomial naïve bayes dimulai dengan menghitung peluang masing-masing label seperti pada Persamaan 2. Contoh perhitungan menggunakan hasil ekstraksi fitur pada Tabel 5.

$$P(V_{positif}) = \frac{|Docs_j|}{|training|}$$

$$P(V_{positif}) = \frac{2}{3}$$

$$P(V_{positif}) = 0,667$$

$$P(V_{negatif}) = \frac{|Docs_j|}{|training|}$$

$$P(V_{negatif}) = \frac{1}{3}$$

$$P(V_{negatif}) = 0,333$$

Berikut adalah contoh sampel data uji yang nantinya digunakan untuk perhitungan dalam penelitian ini, tercantum pada tabel 6.

Tabel 6. Data Uji

Content
[buruk, belanja, barang]

Setelah itu, data uji yang sudah melalui preprocessing dihitung probabilitas setiap kata terhadap label menggunakan Persamaan 3.

Perhitungan label positif :

$$\begin{aligned} \text{a. } P(\text{buruk, positif}) &= (0+1) / (6+9) \\ &= 1 / 15 \\ &= 0,0667 \end{aligned}$$

$$\begin{aligned} \text{b. } P(\text{belanja, positif}) &= (2+1) / (6+9) \\ &= 3 / 15 \\ &= 0,2 \end{aligned}$$

$$\begin{aligned} \text{c. } P(\text{barang, positif}) &= (0+1) / (6+9) \\ &= 1 / 15 \\ &= 0,0667 \end{aligned}$$

Perhitungan label negatif :

$$\begin{aligned} \text{a. } P(\text{buruk, negatif}) &= (1+1) / (3+9) \\ &= 2 / 12 \\ &= 0,167 \end{aligned}$$

$$\begin{aligned} \text{b. } P(\text{belanja, negatif}) &= (21+1) / (3+9) \\ &= 2 / 12 \\ &= 0,167 \end{aligned}$$

$$\begin{aligned} \text{c. } P(\text{barang, negatif}) &= (0+1) / (3+9) \\ &= 1 / 12 \\ &= 0,0834 \end{aligned}$$

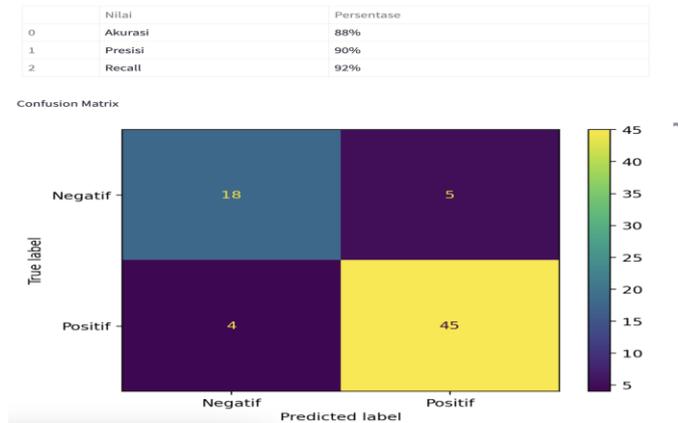
Setelah diperoleh peluang dari setiap label dan peluang kata-kata dalam data uji terhadap label, langkah selanjutnya adalah melakukan klasifikasi menggunakan persamaan 1.

$$\begin{aligned} P(\text{positif, dl}) &= 0,05 \times 0,05 \times 0,05 \times 0,667 \\ &= 0,00008337500000000002 \\ P(\text{negatif, dl}) &= 0,0667 \times 0,1333 \times 0,0667 \times 0,333 \\ &= 0,00019748133332099998 \end{aligned}$$

Dari hasil perhitungan di atas terungkap bahwa kemungkinan suatu dokumen bertemu dengan label negatif lebih tinggi daripada kemungkinan suatu dokumen bertemu dengan label positif. Dari sini dapat disimpulkan bahwa sampel data uji coba diklasifikasikan sebagai negatif.

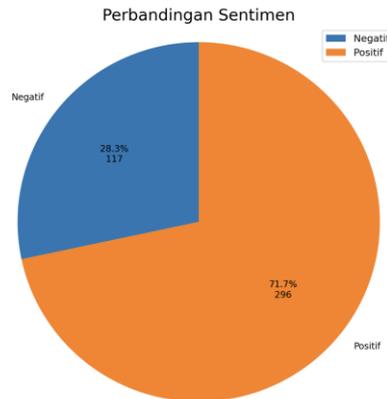
3.7 Tahap Pengujian dan Evaluasi

Pengujian dijalankan menurut rencana pengujian yang tercantum di Sub-Bab 2.9. Pengujian dikerjakan secara langsung oleh sistem yang sudah dibuat. Hasil pengujian pada data pertama bisa dilihat pada Gambar 2.



Gambar 2. Hasil Pengujian

Didapatkan hasil pengujian terhadap dataset bahwa nilai akurasi, presisi, dan recall masing-masing sebesar 88%, 90%, dan 92%. Selain itu persentase sentimen pada dataset juga menunjukkan bahwa sentimen positif mempunyai persentase yang lebih tinggi daripada sentimen negatif, persentase sentimen dapat dilihat pada Gambar 3.



Gambar 3. Persentase Sentimen Dataset

4. KESIMPULAN

Dari hasil penelitian yang telah dilakukan, dari dataset yang berjumlah 413 data bahwa arah pandang atau sentimen pengguna yang menggunakan aplikasi wahyoo mendapatkan 296 ulasan positif sebesar (71,7%) dan 117 ulasan negatif (28,3%). Pada penelitian ini tahap preprocessing yang efektif sangat menentukan hasil yang optimal pada tahap berikutnya. Penerapan metode ekstraksi fitur CountVectorizer dan Algoritme Multinomial Naïve Bayes dalam melakukan analisis sentimen mampu berjalan dengan efektif, dengan hasil nilai pengujian dan evaluasi yaitu akurasi 88%, presisi 90%, recall 92%.

Sistem analisis sentimen ulasan pengguna aplikasi wahyoo diharapkan dapat menambahkan daftar kosakata slang atau bahasa gaul, singkatan atau tidak formal, dan stop word dengan variasi kata atau gaya penulisan pada ulasan yang akan diolah, dapat menggunakan database sebagai media penyimpanan dan juga dapat menambahkan fitur import file excel.

Penelitian berikutnya diharapkan dapat menggabungkan lebih banyak teknik dalam analisis data dan penyelesaian masalah, sehingga diperoleh sebuah model sistem analisis sentimen yang lebih efisien dan efektif dalam pengolahan/penyajian informasi.

DAFTAR PUSTAKA

- [1] F. V. Sari dan A. Wibowo, “Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naïve Bayes Classifier Berbasis Konversi Ikon Emosi,” *J. SIMETRIS*, vol. 10, no. 2, hal. 681–686, 2019.
- [2] S. N. J. Fitriyyah, N. Safriadi, dan E. E. Pratama, “Analisis Sentimen Calon Presiden Indonesia 2019 dari Media Sosial Twitter Menggunakan Metode Naive Bayes,” *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 3, hal. 279, 2019, doi: 10.26418/jp.v5i3.34368.
- [3] T. A. Mutiara, Y. Alkhalifi, N. A. Mayangky, dan W. Gata, “Analisis Sentimen Opini Publik Mengenai Larangan Mudik pada Twitter Menggunakan Naïve Bayes,” *J. CoreIT J. Has. Penelit. Ilmu Komput. dan Teknol. Inf.*, vol. 6, no. 2, hal. 85, 2020, doi: 10.24014/coreit.v6i2.9727.
- [4] V. Rizya *et al.*, “PADA GOOGLE PLAY STORE DENGAN MULTINOMIAL NAÏVE SENTIMENT ANALYSIS OF IPUSNAS APPLICATION USER REVIEWS ON GOOGLE PLAY STORE WITH MULTINOMIAL NAÏVE,” vol. 2, no. April, hal. 182–190, 2023.
- [5] S. Khairunnisa, A. Adiwijaya, dan S. Al Faraby, “Pengaruh Text Preprocessing terhadap Analisis Sentimen Komentar Masyarakat pada Media Sosial Twitter (Studi Kasus Pandemi COVID-19),” *J. Media Inform. Budidarma*, vol. 5, no. 2, hal. 406–414, 2021, doi: 10.30865/mib.v5i2.2835.
- [6] A. R. Alita Debby, “Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier,” *J. Komputasi*, vol. 8, no. 2, hal. 50–58, 2020, doi: 10.23960/komputasi.v8i2.2615.
- [7] O. I. Gifari, M. Adha, F. Freddy, dan F. F. S. Durrand, “Analisis Sentimen Review Film Menggunakan TF-IDF dan Support Vector Machine,” *J. Inf. Technol.*, vol. 2, no. 1, hal. 36–40, 2022, doi: 10.46229/jifotech.v2i1.330.
- [8] D. Darwis, N. Siskawati, dan Z. Abidin, “Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional,” *J. Tekno Kompak*, vol. 15, no. 1, hal. 131–145, 2021, doi: 10.33365/jtk.v15i1.744.
- [9] D. Normawati dan S. A. Prayogi, “Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter,” *J. Sains Komput. Inform.*, vol. 5, no. 2, hal. 697–711, 2021, [Daring]. Tersedia pada: <http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/369>
- [10] J. Kalyzta, M. A. Willdan, S. Halfiani, dan I. Indra, “Penerapan Analisis Sentimen Ujaran Kebencian Terhadap Vaksinasi Covid-19 Pada Tweet Berbahasa Indonesia Menggunakan Algoritme K-Nearest Neighbor,” *IDEALIS Indonesia J. Inf. Syst.*, vol. 5, no. 2, hal. 87–97, 2022, doi: 10.36080/idealis.v5i2.2959.