

Analisis Sentimen Keberhasilan Aplikasi *Meet You* Dalam Memprediksi Siklus Menstruasi Menggunakan Metode *Naive Bayes*

Novia Elma Hidayah^{1,*}, Nur Hidayah Setia Ningrum², Novita Rahmayuna³, Nutriana Hidayati⁴

^{1,2,4}Fakultas Teknologi Informasi dan Komunikasi, Sistem Informasi, Universitas Semarang, Semarang, Indonesia

³*School of Information Systems*, Sistem Informasi, Universitas Bina Nusantara, Kota Semarang

Email: ^{1,*}elmahidayah05@email.com, ²dayahstym@gmail.com, ³novita.rahmayuna@binus.ac.id,

⁴anna@usm.ac.id

*) *Email* Penulis Utama

Abstrak– Aplikasi pemantau siklus menstruasi semakin banyak diminati di zaman digitalisasi kesehatan reproduksi. Aplikasi *MeetYou* merupakan salah satu aplikasi yang sering digunakan untuk meramalkan siklus menstruasi. Namun, masih sedikit penelitian yang mengevaluasi efektivitas aplikasi itu berdasarkan perasaan pengguna. Penelitian ini bertujuan untuk mempelajari respon pengguna mengenai efektivitas aplikasi *MeetYou* dalam memperkirakan siklus menstruasi dengan memanfaatkan algoritma *Naive Bayes*. Data penelitian dikumpulkan melalui *web scraping* dari ulasan pengguna aplikasi *MeetYou* di *Google Play Store* yang diperoleh dari 26 April 2024 sampai 17 Mei 2025 sejumlah 1000 data. Metode penelitian mencakup langkah-langkah pengumpulan data, pemilihan, pra-pemrosesan (pembersihan, pengubahan huruf besar kecil, tokenisasi, pemfilteran), transformasi, penambangan data menggunakan algoritma *Naive Bayes Multinomial*, dengan penilaian memakai *confusion matrix* untuk menilai akurasi, presisi, dan recall. Algoritma *Naive Bayes* dipilih karena kesederhanaan, efisiensi, dan kemampuannya yang unggul dalam mengklasifikasikan teks pada set data yang besar. Hasil penelitian mengungkapkan bahwa algoritma *Naive Bayes* berhasil dalam menganalisis sentimen dari ulasan pengguna aplikasi *MeetYou* dengan persentase ketepatan 80.67%. Data riset dibagi menjadi 80% data pelatihan dan 20% data pengujian, sehingga dapat memberikan masukan penting bagi pengembang untuk menambah kualitas layanan prediksi siklus menstruasi di waktu yang akan datang

Kata Kunci: Menstruasi, *Meet You*, *Naive Bayes*, Analisis Sentimen, Prediksi

Abstract– Menstrual cycle monitoring applications are increasingly popular in the era of digital reproductive health. The *MeetYou* application is one application frequently used to predict menstrual cycles. However, there is still little research evaluating the effectiveness of the application based on user sentiment. This study aims to study user responses regarding the effectiveness of the *MeetYou* application in predicting menstrual cycles using the *Naive Bayes* algorithm. Research data was collected through web scraping from 1000 *MeetYou* application user reviews on the *Google Play Store* obtained from April 26, 2024 to May 17, 2025. The research method includes data collection, selection, pre-processing (cleaning, case conversion, tokenization, filtering), transformation, data mining using the *Naive Bayes Multinomial* algorithm, with an assessment using a *confusion matrix* to assess accuracy, precision, and recall. The *Naive Bayes* algorithm was chosen because of its simplicity, efficiency, and superior ability to classify text on large datasets. The results revealed that the *Naive Bayes* algorithm successfully analyzed sentiment from *MeetYou* application user reviews with an accuracy rate of 80.67%. The research data is divided into 80% training data and 20% testing data, so that it can provide important input for developers to improve the quality of menstrual cycle prediction services in the future.

Keywords: Menstruation, *Meet You*, *Naive Bayes*, Sentiment Analysis, Prediction

1. PENDAHULUAN

Pertumbuhan teknologi yang cepat di Indonesia membawa banyak perubahan di hampir semua bidang kehidupan [1]. Salah satunya di bidang kesehatan telah melahirkan berbagai inovasi aplikasi mobile. Aplikasi-aplikasi tersebut hadir untuk membantu pengguna dalam memantau kondisi fisik maupun siklus biologis. Salah satu jenis aplikasi yang semakin populer adalah pencatatan siklus menstruasi, seperti *MeetYou*. Aplikasi ini dirancang untuk membantu wanita dalam memantau kesehatan reproduksi melalui fitur prediksi menstruasi, ovulasi, serta pencatatan gejala harian. Keberadaan aplikasi ini sejalan dengan meningkatnya kesadaran akan pentingnya pengelolaan kesehatan reproduksi di era digital. Selain memberikan informasi mengenai pola siklus menstruasi, aplikasi ini juga berfungsi sebagai pendukung dalam perencanaan aktivitas harian dan pemantauan kondisi tubuh. Salah satu faktor yang mempengaruhi pola siklus menstruasi adalah stress [2]. Namun, meskipun popularitas aplikasi pencatatan siklus cukup tinggi, kajian ilmiah yang menilai tingkat akurasi prediksi dan kenyamanan penggunaannya masih terbatas. Penilaian keberhasilan aplikasi sering kali hanya menekankan aspek teknis atau

medis, sementara persepsi dan sentimen pengguna sebagai representasi kepuasan serta harapan masih jarang diteliti. Karena itu, diperlukan pendekatan dengan menganalisis perasaan pengguna dalam ulasan mereka, yang bisa memberikan gambaran lebih baik bagi pengembang agar bisa meningkatkan kualitas aplikasi.

Penilaian terhadap keberhasilan aplikasi sering kali hanya berfokus pada aspek teknis atau medis, bukan dari persepsi dan sentimen pengguna yang sebenarnya. Oleh karena itu, dibutuhkan pendekatan berbasis analisis sentimen terhadap ulasan pengguna sebagai representasi langsung dari kepuasan dan harapan mereka. Penelitian ini penting untuk membantu pengembang aplikasi memahami kebutuhan pengguna secara lebih menyeluruh serta meningkatkan kualitas layanan berdasarkan data real-time yang aktual dan masif.

Penelitian ini dilatarbelakangi oleh kurangnya kajian yang secara spesifik meneliti persepsi pengguna terhadap aplikasi pencatatan siklus menstruasi, khususnya *MeetYou*, melalui pendekatan analisis sentimen.. Kebanyakan studi penelitian yang dilakukan hanya mengkaji prediksi siklus berdasarkan data biologis, bukan dari sudut pandang pengguna. Adapun Penelitian terdahulu terkait aplikasi pelacak menstruasi antara lain dilakukan oleh Damayanti et al. (2025), yang menganalisis aplikasi K24Klik dan memperoleh hasil bahwa Naïve Bayes mampu mengklasifikasikan sentimen pengguna dengan akurasi lebih dari 80% [3]. Penelitian lain oleh Insan et al. (2023) menggunakan pendekatan serupa pada aplikasi BRImo, dan juga menunjukkan efektivitas pendekatan ini dalam menangkap persepsi pengguna secara akurat melalui ulasan *PlayStore* [4]. Aplikasi ini sudah di gunakan oleh 10 juta orang dengan 177.000 ulasan sentiment.

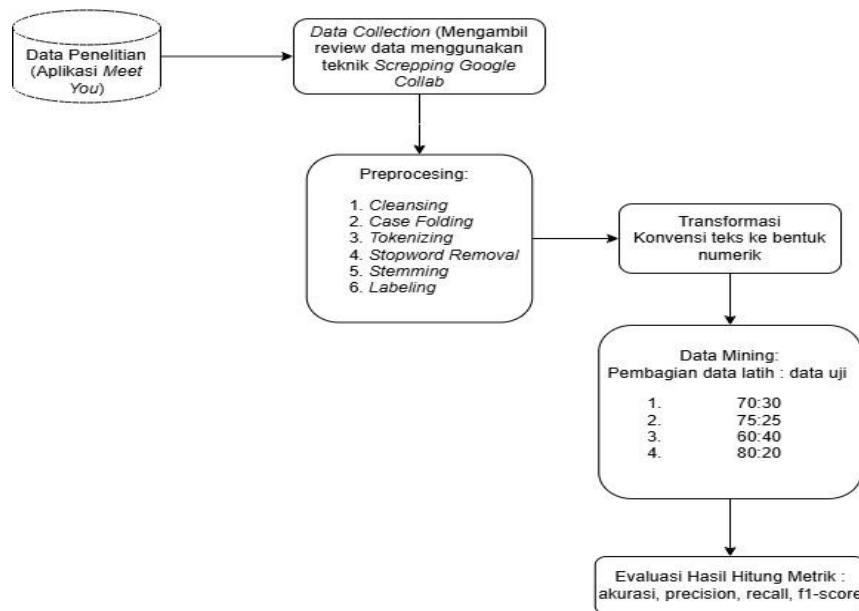
Meskipun aplikasi ini sudah banyak digunakan, cara mengevaluasinya berdasarkan pengalaman langsung pengguna masih terbatas. Oleh karena itu, penelitian ini fokus pada tiga masalah utama. Pertama, bagaimana persepsi pengguna terhadap aplikasi *MeetYou* dalam memprediksi siklus menstruasi berdasarkan ulasan di Google Play Store. Kedua, berapa besar efektifitas algoritma Naïve Bayes dalam mengklasifikasikan perasaan puas dalam menggunakan aplikasi dari tersebut. Ketiga, apa saja saran yang bisa diberikan kepada pengembang aplikasi berdasarkan hasil analisis sentimen yang diperoleh..

Algoritma Naïve Bayes adalah salah satu algoritma yang cukup populer, karena sifatnya yang sederhana dan kemampuannya dalam mengklasifikasikan teks dengan tingkat keakuratan yang memadai. Algoritma ini bekerja berdasarkan prinsip probabilitas Bayes dengan asumsi bahwa setiap karakteristik dalam data bersifat independent [5]. Naïve Bayes sangat cocok digunakan untuk membantu dalam pengambilan keputusan [6]. Algoritma ini juga bisa dipertimbangkan sebagai pilihan yang tepat dalam menganalisis komentar pengguna di Google Play Store karena mudah diterapkan dan mudah dipahami. Untuk meneliti dan menganalisis fenomena tersebut, dibutuhkan metode klasifikasi yang mampu mengelompokkan komentar pengguna ke dalam beberapa kategori, dalam penelitian ini dibagi menjadi kategori positif dan negatif [7]. Namun, terdapat masalah ketika ulasan pengguna memiliki konteks yang beragam, sehingga mengakibatkan perbedaan tingkat akurasi klasifikasi[7].

Tujuan dari penelitian ini untuk mengetahui efektifitas Algoritma *Naive Bayes* dalam menganalisis sentimen pengguna menggunakan data ulasan pada aplikasi *Meetyou* di *Google Play Store*.

2. METODE PENELITIAN

Penelitian ini menggunakan algoritma *Naive Bayes* untuk menganalisis sentimen dari ulasan aplikasi *Meetyou*. Dengan Tahapan awal pengumpulan data (*Data collection*), *pre-processing* data, *transformasi* data, *Data Mining*, Dan *Evaluation*. Tahapan diagram alur dapat di lihat di gambar 1.



Gambar 1. Diagram Alur Metode Penelitian
Sumber: hasil olahan peneliti (2025)

2.1 Data Collection

Pengumpulan data atau pemilihan data dilakukan melalui *Google Play Store* dengan memanfaatkan teknik *web scraping*. Metode *web scraping* merupakan suatu pendekatan untuk memperoleh informasi secara otomatis dari situs web, di mana data diperoleh dengan cara mengekstraksi konten yang relevan sesuai dengan kebutuhan penelitian [8]. Dengan demikian, data yang diperoleh melalui metode ini dianggap lebih representatif dan memberikan gambaran yang akurat terkait sentimen pengguna terhadap aplikasi yang diteliti [9]. Pengumpulan data dilakukan pada rentang waktu 26 April 2024 hingga 17 Mei 2025, dengan sumber data berasal dari ulasan pada aplikasi *Meetyou* yang ada di platform *Google Play Store* dengan jumlah 1000 data. Aplikasi ini adalah perangkat lunak yang dipakai untuk membantu perempuan dalam memprediksi siklus menstruasi setiap bulannya.

2.2 Pre-processing Data

Pre-processing adalah proses penting dan proses eksplorasi data [10]. Pada Saat mengumpulkan data, yang didapat biasanya merupakan data yang tidak terstruktur dan mengandung banyak karakter. Tujuan dari *pre-processing* adalah untuk menghilangkan *noise* tersebut. Tahap *preprocessing text* dalam penelitian ini meliputi 5 tahapan utama, yaitu *cleansing*, *case folding*, *tokenizing*, *stopwatch removal*, *stemming*, dan terakhir proses *filtering*. Sehingga dari proses *pre-processing* di hasilkan data yang sudah di bersihkan menjadi 746 data dari 1000 data.

2.3 Transformation

Pada tahap ini dilakukan proses penimbangan kata (TF), yaitu pemberian nilai bobot pada setiap kata. Nilai bobot tersebut menunjukkan seberapa penting suatu kata dalam sebuah dokumen [11]. Data yang telah disediakan akan dikonversi menjadi bentuk numerik dengan menggunakan teknik TF-IDF (Frekuensi Istilah - Frekuensi Balik Dokumen), sehingga bisa dibaca dan diproses oleh algoritma klasifikasi.

2.4 Data Mining

Penggalian data adalah sebuah proses analisis yang bertujuan untuk mengambil pola, pengetahuan, atau informasi yang signifikan dari sekumpulan data yang besar. Prosesnya melibatkan analisis teks untuk mendapatkan informasi yang relevan sesuai kebutuhan tertentu. Berbeda dengan data terstruktur dalam database, teks memiliki sifat tidak terstruktur, ambigu, dan sulit diproses. Meskipun demikian, dalam era modern, teks menjadi salah satu cara umum dalam pertukaran informasi secara resmi [12]. Proses klasifikasi dalam data mining pada penelitian ini dilakukan dengan memakai algoritma *Naive Bayes Multinomial* dalam pemodelan probabilistik berbasis Teorema Bayes, dengan asumsi bahwa setiap fitur (misalnya kata dalam dokumen) bersifat independen satu sama lain, yang dikenal dengan istilah *naive assumption*), serta fitur tersebut

mengikuti distribusi multinomial. Model ini menghitung probabilitas dari sebuah dokumen ddd termasuk dalam kelas ccc berdasarkan frekuensi kata yang muncul dalam dokumen. Persamaan (1) menunjukkan rumus perhitungan dari *Naive Bayes Multinomial* :

$$P(C_k | x) \propto P(C_k) \cdot \prod P(w_i | C_k)^{x_i} \quad (1)$$

Keterangan :

C_k = kelas ke-k (misalnya: positif, negatif, spam, not spam)

x = (x_1, x_2, \dots, x_n) : representasi dokumen dalam bentuk vektor jumlah kata

w_i = kata ke-i dalam kosakata

x_i = jumlah kata w_i dalam dokumen

$P(C_k)$ = probabilitas awal (prior) dari kelas C_k

$P(w_i | C_k)$ = probabilitas kata w_i muncul dalam dokumen dari kelas C_k

2.5 Evaluation

Selanjutnya tahap evaluasi, melakukan metrik evaluasi seperti precision, recall, f1-score, dan dukungan untuk setiap kategori (negatif dan positif)[13]. Di tahap ini pengembang aplikasi akan mendapat informasi pengembangan apa yang harus dilakukan agar aplikasinya dapat memenuhi kebutuhan pemakai.

3. HASIL DAN PEMBAHASAN

Hasil dan Analisis ini membahas mengenai analisis sentimen keberhasilan penggunaan aplikasi *Meetyou* dalam memprediksi siklus menstruasi menggunakan metode *Naive Bayes Multinomial* dengan data yang diolah menggunakan *Google Collab* dengan melalui beberapa tahapan yaitu tahapan data *selection*, *preprocessing*, *transformasi*, *data mining*, dan *evaluasi*.

3.1 Data Collection

Diproses data *collection* ini di ambil data menggunakan Teknik *web scraping* dengan data yang dikumpulkan mulai tanggal 26 April 2024 sampai 17 Mei 2025 sehingga menghasilkan data sejumlah 1000 data mentah. Pada data *collection* yang di hasilkan ini masih berupa data asli dimana masih banyak noise di dalamnya. Berikut di sajikan data mentahnya pada Tabel 1:

Tabel 1. Data Hasil *Scraping* menggunakan *Google Collab*, Sumber: hasil olahan peneliti (2025)

<i>Score</i>	<i>Content</i>
1	Terlalu banyak iklan δÿ·Žδÿ·»
2	Terlalu banyak iklan
4	Bagus
5	mantap sekali
...1000	

3.2 Pre-processing

Di tahap berikutnya ada *pre-processing* dimana di tahap ini ada beberapa langkah yang harus dilakukan yaitu data *cleaning*, *case folding*, *tokenizing*, dan *filtering* atau labeling. Dari tahap *pre-processing* dan *transformation* di dapat data yang sudah bersih dengan jumlah data 746 dari 1000 data yang sudah siap di proses ke tahap

selanjutnya. Disajikan hasil tahapan *pre-processing* terhadap data ulasan pengguna, mulai dari teks asli hingga proses *stemming* pada Tabel 2.

Tabel 2. Hasil tahapan *Pre-processing* data dari teks asli sampai proses *filtering*, Sumber: hasil olahan peneliti (2025)

Tahapan <i>Pre-processing</i>	Hasil
Data Asli	Saya suka menggunakan aplikasi ini
<i>Cleansing</i>	Saya suka menggunakan aplikasi ini
<i>Case Folding</i>	saya suka menggunakan aplikasi ini
<i>Tokenizing</i>	['saya', 'suka', 'menggunakan', 'aplikasi', 'ini']
<i>Stopword Removal</i>	['suka', 'menggunakan', 'aplikasi']
<i>Stemming</i>	['suka', 'guna', 'aplikasi']

Proses Selanjutnya dilakukan pelabelan secara otomatis berdasarkan nilai skor yang terdapat dalam data. Skor dengan nilai ≤ 2 dikategorikan sebagai sentimen 'negatif', sedangkan skor > 2 dikategorikan sebagai sentimen 'positif'. Hasil pelabelan ini disimpan dalam kolom 'label' dan digunakan sebagai target dalam pelatihan model klasifikasi *Naive Bayes Multinomial*. Sehingga pengembang aplikasi dapat melihat sentiment pengguna lebih condong ke positif atau ke negatif. Diberikan ilustrasi hasil penandaan otomatis untuk ulasan pengguna aplikasi *MeetYou* yang terbagi dua jenis perasaan, yaitu positif dan negatif pada Tabel 3.

Tabel 3. Source Code Python Proses Pelabelan, Sumber: hasil olahan peneliti (2025)

No	Hasil <i>Pre-processing</i> Data	Sentiment
1	Kenapa sekarang iklanya banyak lama sekali yaa	Negatif
2	kenapa gak bisa di download saya udah berulang ulang men download tapi tetap gak bisa ngabisin data aja	Negatif
3	Saya suka menggunakan aplikasi ini	Positif
4	Banyak tepat perhitungannya	Positif

3.3 Transformation

Tahap ini, dilakukan pengolahan teks dengan menggunakan metode TF-IDF (Frekuensi Kata - Frekuensi Terbalik Dokumen). *Transformasi* ini bertujuan merubah data kalimat mentah menjadi representasi data angka menggunakan bobot kata (fitur) yang merefleksikan tingkat signifikansi kata dalam sebuah dokumen terhadap keseluruhan korpus. Proses transformasi dilakukan dengan memanfaatkan *TfidfVectorizer* dari pustaka *scikit-learn*. Sebagai contoh, digunakan 3 data sampel ulasan yang sudah dibersihkan pada Tabel 4:

Tabel 4. Sampel ulasan yang sudah melewati tahapan *pre-processing*, Sumber: hasil olahan peneliti (2025)

Variabel	Ulasan
D1	Sukak sekali aplikasinya

D2	Lumayan jadi ada aplikasi pengingat haid
D3	kurangi iklan

Dari 3 sampel ulasan yang sudah melewati tahapan *pre-processing* kemudian dilakukan tahapan *tokenizing* yaitu Mengurai teks menjadi kata-kata (istilah) yang ada pada Tabel 5.

Tabel 5.Tahapan *tokenizing*, Sumber: hasil olahan peneliti (2025)

<i>Variabel</i>	<i>Tokenizing</i>
D1	['sakak', 'sekali', 'aplikasinya']
D2	['lumayan', 'jadi', 'ada', 'aplikasi', 'pengingat', 'haid']
D3	['kurangi', 'iklan']

Setelah melalui proses *tokenizing* selanjutnya dilakukan tahapan *Stopword Removal* yaitu menghilangkan kata umum (seperti "jadi", "ada", dll) yang tidak berarti secara signifikan. Berikut di sajikan tabel penghapusan pada Tabel 6.

Tabel 6.Tahapan *Stopword Removal*, Sumber: hasil olahan peneliti (2025)

<i>Variabel</i>	<i>Stopword Removal</i>
D1	['sakak', 'sekali', 'aplikasinya']
D2	['lumayan', 'aplikasi', 'pengingat', 'haid']
D3	['kurangi', 'iklan']

Pada tahap *transformasi* teks, salah satu langkah penting adalah menentukan kata unik (*vocabulary*) dari seluruh dokumen dalam korpus. Kata unik adalah sekumpulan istilah yang berbeda dan hanya dihitung satu kali, meskipun muncul berulang kali di berbagai dokumen. Dengan cara ini, setiap dokumen dapat direpresentasikan dalam bentuk vektor sesuai dengan keberadaan kata-kata unik tersebut. Dari 3 sample *pre-prpcesing* tersebut di peroleh kata unik “aplikasi”.

Kata "aplikasi" kemudian digunakan untuk menghitung frekuensi kemunculan suatu kata di setiap dokumen *TF(Frekuensi Term)*. Penghitungan ini menggunakan Persamaan (2) rumus *Term Frequency (TF)* dan Table 7 hasil *TF (Frekuensi Term)*:

$$TF(t,d) = \frac{\text{jumlah kemunculan kata } t \text{ dalam dokumen } d}{\text{total kata dalam } d} \quad (2)$$

Tabel 7. Hasil *TF (Frekuensi Term)*, Sumber: hasil olahan peneliti (2025)

<i>Variabel</i>	<i>Hasil TF (Frekuensi Term)</i>	<i>Bentuk Desimal</i>
D1	1/3	0,33

D2	1/4	0,25
D3	0/2	0

Langkah selanjutnya menghitung yaitu jumlah dokumen yang mengandung kata tersebut *DF* (*Document Frequency*). Dari ke-3 sampel yang digunakan, kata “aplikasi” muncul pada variabel D1 dan D2. Sehingga di peroleh nilai $DF = 2$. Setelah di peroleh nilai *DF*, dilakukan penghitungan *IDF* (*Inverse Document Frequency*) yaitu menghitung nilai aslogaritma dari rasio total dokumen dengan jumlah dokumen yang mengandung kata itu. Penghitungan ini menggunakan Persamaan (3) rumus *IDF* (*Inverse Document Frequency*):

$$IDF(t) = \log\left(\frac{N}{DF(t)}\right) \quad (3)$$

dengan N adalah jumlah seluruh dokumen. Pada kasus ini:

Jumlah dokumen (N) = 3

Document Frequency (*DF*) kata “aplikasi” = 2

Sehingga di dapatkan penghitungan rumus (3) sebagai berikut :

$$IDF(\text{"aplikasi"}) = \log(2/3) = \log(1,5) \quad (3)$$

Kemudian jika menggunakan algoritma basis 10 di peroleh hasil penghitung persamaan (4):

$$IDF(\text{"aplikasi"}) \approx 0,176 \quad (4)$$

Setelah memperoleh nilai *IDF* (*Inverse Document Frequency*) maka menghitung nilai bobot *TF-IDF* menggunakan penghitungan persamaan (5) rumus *TF-IDF* dan disajikan hasil nilai bobot *TF-IDF* pada Tabel 8 :

$$TF-IDF(t,d) = TF(t,d) \times IDF(t) \quad (5)$$

Keterangan:

$TF(t,d)$ = frekuensi relatif kata t dalam dokumen d

$DF(t)$ = inverse document frequency kata t terhadap seluruh korpus

Tabel 8. Hasil nilai bobot *TF-IDF*, Sumber: hasil olahan peneliti (2025)

Variabel	Hasil TF (aplikasi)	IDF	TF-IDF
D1	1/3	0,176	0.058
D2	1/4	0,176	0,044
D3	0/2	0	0

Hasil dari transformasi Kata “aplikasi” dapat dikategorikan sebagai kata yang cukup umum karena muncul pada sebagian besar dokumen dalam korpus (D1 dan D2). Nilai *Inverse Document Frequency* (*IDF*) yang diperoleh sebesar 0,176 menunjukkan bahwa kata ini sering digunakan, sehingga bobotnya relatif rendah dalam perhitungan *TF-IDF*. Hal ini mengindikasikan bahwa meskipun kata “aplikasi” berkontribusi terhadap representasi dokumen, tingkat signifikansinya tidak setinggi kata-kata lain yang lebih jarang muncul dan lebih spesifik. Dalam konteks klasifikasi sentimen, keberadaan kata “aplikasi” masih berperan sebagai fitur, tetapi tidak menjadi kata kunci utama yang dapat membedakan sentimen antar dokumen. Sebaliknya, kata-kata dengan frekuensi lebih rendah dan konteks yang lebih unik cenderung memiliki bobot *TF-IDF* lebih tinggi, sehingga lebih berpengaruh terhadap hasil klasifikasi.

3.4 Data Mining

Pada tahap ini, proses klasifikasi dilakukan dengan menggunakan algoritma Naïve Bayes Multinomial untuk mengelompokkan data ke dalam masing-masing kategori sentimen.. Di dalam pengklasifikasinya melawati beberapa proses yaitu proses *pre-processing* data dimana di dalamnya ada beberapa tahapan yang dilakukan meliputi pembersihan data yang disebut dengan data *cleaning*, *case folding*, *tokenizing*, dan *filtering* atau labeling. Setelah dilakukan tahap *pre-processing* data di lakukan proses pembagian data antara data *training* dan data *testing* dengan beberapa variasi perbandingan data. Tujuan dari variasi ini adalah untuk mengamati pengaruh ukuran data pelatihan terhadap akurasi dan ketepatan klasifikasi sentimen Berdasarkan hasil klasifikasi menggunakan metode algoritma *Naive Bayes Multinomial* menghasilkan perhitungan *accuracy*, *precision*, dan *recall*. Perhitungan tersebut dibagi menjadi 4 persentasi data *testing* dan data *training* untuk di bandingkan dan di ambil nilai *accuracy* yang paling tinggi. Percobaan dilakukan dengan empat skenario rasio pembagian data seperti ditunjukkan pada Tabel 9 dan Tabel 10 berikut:

Tabel 9. Rasio Pembagian data training dan testing, Sumber: hasil olahan peneliti (2025)

Percobaan	Training Set (data Latih)	Testing Set (Data Uji)
1	70	30
2	75	25
3	60	40
4	80	20

Tabel 10. Perbandingan Hasil Percobaan, Sumber: hasil olahan peneliti (2025)

Percobaan	Accuracy	Precision (Positif)	Recall (Positif)	F1-Score (Positif)	Precisi on (Negatif)	Recall (Negatif)	F1-Score (Negatif)
1	78.66 %	0.76	0,097	0.85	0.90	0.46	0.61
2	78.19 %	0.75	0.97	0.85	0.91	0.46	0.61
3	78.33 %	0.76	0.96	0.85	0.87	0.45	0.59
4	80.66 %	0.79	0.97	0.87	0.90	0.50	0.64

Model *Naive Bayes Multinomial* dilatih menggunakan data *training* pada masing-masing skenario, kemudian diuji menggunakan data *testing*. Hasil evaluasi dari setiap percobaan akan dibandingkan berdasarkan nilai *accuracy*, *presisi*, dan *recall* untuk mengetahui konfigurasi pembagian data yang menghasilkan performa klasifikasi sentimen paling optimal. Dari percobaan di atas dinyatakan bahwa percobaan ke-4 layak dipilih sebagai model akhir karena memiliki kombinasi terbaik antara akurasi dan keseimbangan kelas. Tantangan utama yang harus diperhatikan adalah meningkatkan *recall* untuk kelas negatif dengan cara menyeimbangkan data atau menggunakan teknik lain seperti *ensemble* dan model *deep learning*. Dalam penelitian ini, hasil analisis setiap kategori sentimen disajikan dalam bentuk *word cloud*. Ukuran besar atau kecil *word cloud* dipengaruhi oleh banyak total kata yang muncul dalam data ulasan. Berikut disajikan gambar 1 *word cloud* positif dan negatif dari percobaan 4.



Gambar 1. Word Cloud Sentimen Positif

Pada gambar diatas merupakan word cloud dari data ulasan Meetyou positif. Terdapat kata-kata yang sering muncul pada ulasan positif adalah “akurat”, “bantu”, “bagus”, dan lain-lain.

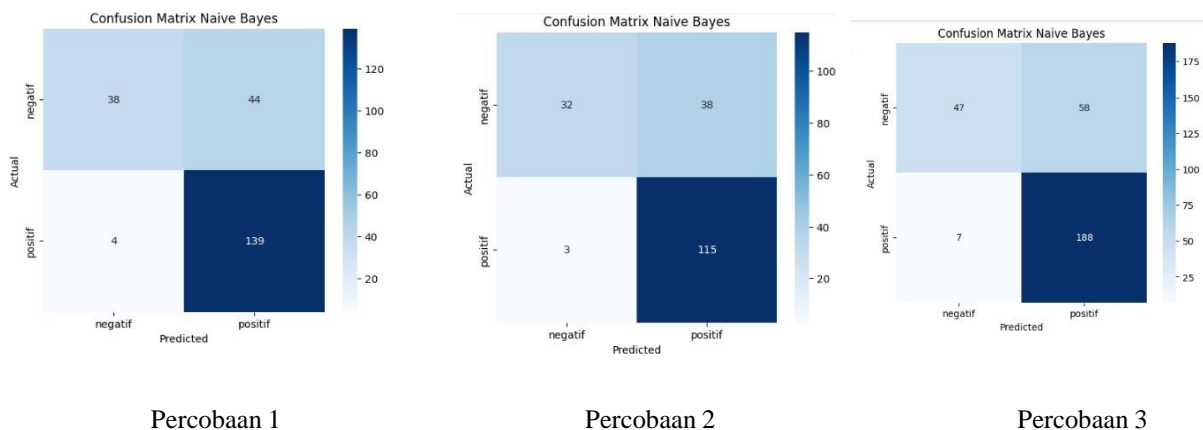


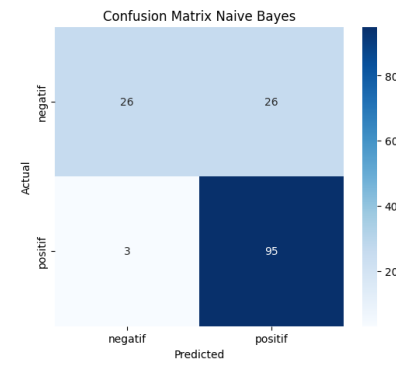
Gambar 2. Word Cloud Sentimen Negatif

Pada gambar diatas merupakan word cloud dari data ulasan Meetyou negatif. Terdapat kata-kata yang sering muncul pada ulasan negatif adalah “banyak iklan”, “jelek”, “kurang”, dan lain-lain.

3.5 Evaluation

Evaluation atau evaluasi untuk mengecek kebenaran hasil klasifikasi dari algoritma *Naive Bayes Multinomial*. Pada tahap evaluasi ini, menggunakan metode *Confusion Matrix* untuk mengukur kinerja dari klasifikasi *Naive Bayes Multinomial* dalam penelitian. Metode *Confusion matrix* menunjukkan jumlah prediksi yang benar dan salah yang dilakukan oleh model, dibandingkan dengan label aktual. Berikut di sajikan gambar 3 *Confusion Matrix* dari percobaan 1 sampai ke 4.





Percobaan 4

Gambar 3. *Confusion Matrix* dari percobaan 1 sampai ke 4

Dan dapat disimpulkan bahwa dari gambar *Confusion Matrix* percobaan 1 sampai 4 yang menghasilkan data terbaik ada pada percobaan data ke 4 dengan perbandingan rasio 80 :20. Berikut di sajikan table 11 *Confusion Matrix* hasil percobaan 4

Tabel 11. *Confusion Matrix* Hasil Percobaan ke-4 (Terbaik), Sumber: hasil olahan peneliti (2025)

	Prediksi Negatif	Prediksi Positif
Aktual Negatif	26	26
Aktual Positif	3	95

Dari *confusion matrix* tersebut, dapat dihitung metrik evaluasi sebagai berikut:

$$\begin{aligned}
 Accuracy &= (TP + TN) / (TP + TN + FP + FN) \\
 &= (95 + 26) / (95 + 26 + 26 + 3) \\
 &= 121 / 150 \\
 &= 80.67\%
 \end{aligned}$$

$$\begin{aligned}
 Precision \text{ (Positif)} &= TP / (TP + FP) \\
 &= 95 / (95 + 26) \\
 &= 95 / 121 \\
 &= 0.79
 \end{aligned}$$

$$\begin{aligned}
 Recall \text{ (Positif)} &= TP / (TP + FN) \\
 &= 95 / (95 + 3) \\
 &= 95 / 98 \\
 &= 0.97
 \end{aligned}$$

$$\begin{aligned}
 F1-Score \text{ (Positif)} &= 2 \times (Precision \times Recall) / (Precision + Recall) \\
 &= 2 \times (0.79 \times 0.97) / (0.79 + 0.97) \\
 &= 0.87
 \end{aligned}$$

$$\begin{aligned}
 Precision \text{ (Negatif)} &= TN / (TN + FN) \\
 &= 26 / (26 + 3) \\
 &= 26 / 29 \\
 &= 0.90
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall (Negatif)} &= \text{TN} / (\text{TN} + \text{FP}) \\
 &= 26 / (26 + 26) \\
 &= 26 / 52 \\
 &= 0.50
 \end{aligned}$$

$$\begin{aligned}
 \text{F1-Score (Negatif)} &= 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \\
 &= 2 \times (0.90 \times 0.50) / (0.90 + 0.50) \\
 &= 0.64
 \end{aligned}$$

Model memiliki akurasi cukup tinggi (80.67%), artinya sebagian besar prediksi sentimen sesuai dengan label sebenarnya. Recall positif tinggi menunjukkan bahwa model sangat baik dalam mengenali data dengan sentimen positif. Namun, recall negatif rendah menunjukkan bahwa masih banyak data dengan sentimen negatif yang tidak dikenali dengan baik oleh model. Precision yang tinggi pada kedua kelas menunjukkan bahwa saat model membuat prediksi, hasilnya cenderung benar.

4. KESIMPULAN

Klasifikasi Naïve Bayes Multinomial berhasil dalam mengklasifikasikan sentimen pengguna terhadap aplikasi Meetyou, dengan tingkat akurasi tertinggi mencapai 80.67%. Ini menunjukkan bahwa metode ini sesuai untuk menganalisis teks ulasan pengguna yang menggunakan bahasa Indonesia. Dari keseluruhan 746 data ulasan yang telah diproses, sebagian besar mengindikasikan sentimen positif, menunjukkan bahwa mayoritas pengguna merasa puas dengan fitur prediksi siklus menstruasi pada aplikasi *MeetYou*. Model menunjukkan precision dan recall yang tinggi untuk kelas positif, tetapi recall untuk kelas negatif masih rendah. Ini menunjukkan bahwa model masih menghadapi kesulitan dalam mengenali ulasan yang memiliki sifat negatif, yang mungkin menjadi indikasi adanya tantangan dalam menyeimbangkan representasi data di antara kelas-kelas. Kata awan dari setiap sentimen berhasil menunjukkan kata-kata utama yang tampak dalam ulasan, seperti “praktis” dan “hebat” untuk ulasan positif, serta “lama” dan “terlalu banyak iklan” untuk ulasan negative. Disarankan untuk melakukan penyeimbangan data (seperti oversampling untuk kelas negatif) agar model lebih seimbang dalam mengolah kedua jenis sentimen. Tujuan dari ini adalah untuk meningkatkan recall kelas negatif. Untuk penelitian selanjutnya, direkomendasikan untuk mencoba algoritma lain seperti Random Forest, SVM dengan bobot kelas, atau metode deep learning (LSTM, CNN) yang dapat menangkap konteks kalimat yang lebih rumit. Penelitian dapat diperluas dengan menambahkan dataset dari sumber lain (seperti App Store, forum kesehatan, atau media sosial) untuk memperkaya variasi bahasa serta ungkapan pengguna. Pengkajian menyeluruh terhadap komentar negatif bisa memberikan umpan balik langsung bagi pengembang, terutama mengenai isu seperti iklan yang berlebihan atau kegagalan aplikasi yang sering dikeluhkan. Analisis sentimen yang diperoleh dapat digunakan sebagai umpan balik terus-menerus untuk pengembangan aplikasi, dengan tujuan meningkatkan akurasi prediksi dan pengalaman pengguna.

UCAPAN TERIMAKASIH

Dengan rasa syukur yang mendalam, kami ingin menyampaikan penghargaan yang sebesar-besarnya kepada setiap individu dan kelompok yang telah memberikan kontribusi penting dalam proses penelitian ini. Ucapan terima kasih yang tulus kami sampaikan kepada dosen pembimbing yang telah memberikan bimbingan, pendidikan, dan arahan yang sangat berharga selama tahap penelitian ini. Kami juga ingin menyampaikan penghargaan yang dalam kepada keluarg tercinta, yang dengan sepenuh hati memberikan kasih sayang, doa, serta dukungan moral yang tanpa henti. Ketulusan dan perhatian kalian menjadi sumber kekuatan yang membuat kami tetap teguh dalam mengatasi berbagai rintangan. Selain itu, kepada teman-teman serta kolega yang telah berbagi energi positif, ide-ide, dan dorongan, kami merasa sangat beruntung atas kehadiran dan bantuan dari kalian semua. Semua bentuk dukungan, perhatian, dan bantuan dari setiap pihak menjadi elemen penting yang memungkinkan penelitian ini diselesaikan dengan baik. Semoga segala kebaikan dan bantuan yang telah diberikan mendapatkan balasan berlipat dari Tuhan Yang Maha Esa.

REFERENCES

- [1] S. A. R. Rizaldi, S. Alam, and I. Kurniawan, "Analisis Sentimen Pengguna Aplikasi JMO (Jamsostek Mobile) Pada Google Play Store Menggunakan Metode Naive Bayes," *STORAGE J. Ilm. Tek. dan Ilmu Komput.*, vol. 2, no. 3, pp. 109–117, 2023, doi: 10.55123/storage.v2i3.2334.
- [2] N. A. Yudita, A. Yanis, and D. Iryani, "Hubungan antara Stres dengan Pola Siklus Menstruasi Mahasiswa Fakultas Kedokteran Universitas Andalas," *J. Kesehat. Andalas*, vol. 6, no. 2, p. 299, 2017, doi: 10.25077/jka.v6i2.695.
- [3] F. Damayanti, A. Rahim, and N. Azmi Verdikha, "Analisis Sentimen Ulasan Pengguna Terhadap Aplikasi K24Klik Di Google Play Store Dengan Algoritma Naive Bayes," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 9, no. 2, pp. 3224–3230, 2025, doi: 10.36040/jati.v9i2.13298.
- [4] M. K. Insan, U. Hayati, and O. Nurdiawan, "Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di," *J. Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 478–483, 2023.
- [5] M. Z. Haq, C. S. Octiva, A. Ayuliana, U. W. Nuryanto, and D. Suryadi, "Algoritma Naive Bayes untuk Mengidentifikasi Hoaks di Media Sosial," *J. Minfo Polgan*, vol. 13, no. 1, pp. 1079–1084, 2024, doi: 10.33395/jmp.v13i1.13937.
- [6] R. Y. Hayuningtyas, "Penerapan Algoritma Naive Bayes untuk Rekomendasi Pakaian Wanita," *J. Inform.*, vol. 6, no. 1, pp. 18–22, 2019, doi: 10.31294/ji.v6i1.4685.
- [7] M. F. El Firdaus, N. Nurfaizah, and S. Sarmini, "Analisis Sentimen Tokopedia Pada Ulasan di Google Playstore Menggunakan Algoritma Naive Bayes Classifier dan K-Nearest Neighbor," *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 5, p. 1329, 2022, doi: 10.30865/jurikom.v9i5.4774.
- [8] T. Rizaldi and H. A. Putranto, "Perbandingan Metode Web Scraping Menggunakan CSS Selector dan Xpath Selector," *Teknika*, vol. 6, no. 1, pp. 43–46, 2017, doi: 10.34148/teknika.v6i1.56.
- [9] H. Junianto, P. Arsi, B. A. Kusuma, and D. I. S. Saputra, "Evaluasi Aplikasi Raileo Melalui Analisis Sentimen Ulasan Playstore Dengan Metode Naive Bayes," *SINTECH (Science Inf. Technol. J.)*, vol. 7, no. 1, pp. 27–40, 2024, doi: 10.31598/sintechjournal.v7i1.1505.
- [10] S. M. Salsabila, A. Alim Murtopo, and N. Fadhilah, "Analisis Sentimen Pelanggan Tokopedia Menggunakan Metode Naive Bayes Classifier," *J. Minfo Polgan*, vol. 11, no. 2, pp. 30–35, 2022, doi: 10.33395/jmp.v11i2.11640.
- [11] Gilbert, Syariful Alam, and M. Imam Sulisty, "Analisis Sentimen Berdasarkan Ulasan Pengguna Aplikasi MyPertamina Pada Google Playstore Menggunakan Metode Naive Bayes," *STORAGE J. Ilm. Tek. dan Ilmu Komput.*, vol. 2, no. 3, pp. 100–108, 2023, doi: 10.55123/storage.v2i3.2333.
- [12] K. Kevin, M. Enjeli, and A. Wijaya, "Analisis Sentimen Penggunaan Aplikasi Kinemaster Menggunakan Metode Naive Bayes," *J. Ilm. Comput. Sci.*, vol. 2, no. 2, pp. 89–98, 2024, doi: 10.58602/jics.v2i2.24.
- [13] E. A. Putra, S. Alam, and I. Kurniawan, "Analisis Sentimen Pengguna MY JNE Menggunakan Algoritma Naive Bayes," *J. Teknol. Inform. dan Komput.*, vol. 10, no. 2, pp. 650–666, 2024, doi: 10.37012/jtik.v10i2.2251.