

# ANALISIS SENTIMEN PADA ULASAN APLIKASI *MOBILE BANKING* MENGGUNAKAN METODE *NAÏVE BAYES* DENGAN KAMUS INSET

Alyaa Nadira<sup>1,\*</sup>, Nanang Yudi Setiawan<sup>2</sup>, Welly Purnomo<sup>3</sup>

<sup>1, 2, 3</sup>Sistem Informasi, Fakultas Ilmu Komputer, Universitas Brawijaya

Jl. Veteran, Ketawanggede, Kec. Lowokwaru, Kota Malang 65145

e-mail: alyaanadira@student.ub.ac.id<sup>1</sup>, nanang@ub.ac.id<sup>2</sup>, wepe@ub.ac.id<sup>3</sup>

\*corresponding author

(Naskah masuk : 10 Januari 2023 Diterima untuk diterbitkan : 29 Maret 2023)

## ABSTRAK

Perkembangan teknologi informasi dan komunikasi berpengaruh pada partisipasi dan interaksi pengguna dalam online platform, seperti contohnya adalah pemberian ulasan pada sebuah aplikasi di Google Play Store. Bagi perusahaan, ulasan dapat digunakan untuk mengetahui kebutuhan pengguna dan dapat digunakan juga sebagai bahan evaluasi untuk mengembangkan dan memperbaiki aplikasinya, namun terkadang dalam sebuah ulasan terdapat ketidaksesuaian antara isi ulasan dengan pemberian rating pada ulasan, yang berarti bahwa sentimen dari sebuah ulasan tidak dapat dilihat dari jumlah rating pada ulasan. Dalam sebuah ulasan juga dapat terjadi kesalahan penulisan baik disengaja ataupun tidak disengaja yang membuat ulasan tersebut susah untuk dibaca dan dimengerti. Pada penelitian ini dilakukan analisis sentimen dengan studi kasus aplikasi Victoria Mobile Banking untuk mengklasifikasikan ulasan aplikasi kedalam ulasan yang bersifat positif atau negatif serta dapat memberi visualisasi hasil analisis ulasan. Proses klasifikasi pada sistem menggunakan metode Naïve Bayes. Data latih yang digunakan adalah data ulasan yang telah diberi label secara otomatis menggunakan kamus InSet yang telah melalui penyesuaian kata dan bobot. Algoritma klasifikasi tersebut diuji menggunakan confusion matrix dan menghasilkan nilai precision 90,4%, recall 100%, f-measure 95%, dan akurasi 93,1%.

**Kata Kunci:** Analisis Sentimen, Naïve Bayes, InSet, Text Mining

## ABSTRACT

Information and communication technologies development has an impact on how users interact and participate on the online platform. Providing feedback or a review of an application in Google Play Store, as an example. From the company point of view, in addition to seeing the opinions of users, reviews could be used to find out the wants and needs of users and could also be a resource for development and improvement for their apps. But occasionally, whether intentionally or not, errors in writing can occur in reviews, thus making it challenging to read, comprehend, and also difficult to classify the review as negative or positive reviews. In addition, there may be a mismatch between a review's content and the rating, which means that the sentiment of the review can't be seen from the rating. In this study, sentiment analysis was carried out with a case study of the Victoria Mobile Banking application to help classify apps reviews as positive or negative reviews and can also provide visualization of the review analysis. The classification process in the system uses the Naïve Bayes method, where the training data used is review data that has been automatically labeled using a modified InSet library. The classification algorithm was evaluated using a confusion matrix, and the results were precision of 90.4%, recall of 100%, f-measure of 95%, and accuracy of 93.1%.

**Keywords:** Sentiment Analysis, Naïve Bayes, InSet, Text Mining

## I. PENDAHULUAN

Teknologi informasi dan komunikasi di zaman ini terus berkembang dengan sangat pesat. Perkembangan tersebut juga berpengaruh pada partisipasi dan interaksi pengguna dalam *online platform*, seperti contohnya adalah pemberian ulasan pada sebuah aplikasi [1], [2]. Ulasan sangat umum di setiap industri, dan banyak orang mengandalkannya untuk mempelajari lebih lanjut tentang sesuatu yang mereka minati, terutama sesuatu yang ingin mereka beli atau gunakan [2].

Pada Google Play Store, pengguna sebuah aplikasi dapat memberikan *rating* (dalam bentuk bintang) dan juga ulasan. *Rating* dan ulasan yang ada pada Google Play Store menggambarkan pendapat dari pengguna aplikasi terhadap aplikasi tersebut. Bagi perusahaan, *rating* dan ulasan (terutama ulasan yang negatif) dapat digunakan untuk mengetahui kebutuhan atau keinginan pengguna dan juga sebagai bahan evaluasi untuk mengembangkan dan memperbaiki aplikasinya [3], [4]. *Rating* dapat mempermudah perusahaan untuk memisahkan antara ulasan yang positif dan ulasan yang negatif, namun ternyata dalam sebuah ulasan terkadang terdapat ketidaksesuaian antara isi ulasan dengan pemberian *rating* pada ulasan, yang berarti bahwa positif atau negatifnya sebuah ulasan tidak dapat dilihat dari *rating* yang diberikan oleh pengguna saja. Contohnya adalah ketika pengguna memberikan ulasan yang positif namun hanya memberikan bintang dua, ataupun sebaliknya. Dalam sebuah ulasan juga dapat terjadi kesalahan penulisan baik disengaja ataupun tidak disengaja yang membuat ulasan tersebut susah untuk dibaca dan dimengerti, sehingga sulit untuk mengklasifikasikan ulasan tersebut kedalam ulasan yang bersifat positif atau negatif.

Analisis sentimen merupakan sebuah teknik *Natural Language Processing* (NLP) yang dapat digunakan untuk mengatasi permasalahan tersebut. Analisis sentimen biasanya digunakan pada data berbentuk teks untuk menentukan apakah suatu data bersifat positif, negatif, atau netral. Tujuan dari dilakukannya analisis sentimen adalah untuk membantu sebuah bisnis atau perusahaan dalam memonitor sentimen dari *customer feedback* terhadap brand itu sendiri ataupun suatu produk, dan juga membantu perusahaan untuk memahami kebutuhan pelanggan dan mengambil keputusan bisnis [5].

Berdasarkan penelitian-penelitian sebelumnya analisis sentimen dapat dilakukan dengan berbagai metode, seperti *K-Nearest Neighbor*, *Support Vector Machines*, *Maximum Entropy*, dan *Naïve Bayes*. Dari beberapa metode tersebut, metode *Naïve Bayes* merupakan metode yang memiliki tingkat akurasi dan kinerja yang lebih baik [6]–[8]. Metode *Naïve Bayes* melakukan klasifikasi data berdasarkan faktor-faktor probabilitas secara efektif dan efisien [9]. Metode *Naïve Bayes* juga memiliki beberapa kelebihan, yaitu metode *Naïve Bayes* memiliki kinerja dan performa yang unggul untuk pengujian pada tipe data kategori serta atribut pada data tidak saling ketergantungan (independen) [10]. Dalam melakukan analisis sentimen, *lexicon-based method* dapat digunakan dalam proses pelabelan sentimen, dan hasilnya dapat digunakan sebagai data latih dalam *Naïve Bayes Classifier* [11].

Berdasarkan latar belakang penelitian dan referensi yang digunakan, penelitian ini akan melakukan analisis sentimen pada studi kasus aplikasi Victoria Mobile Banking. Proses klasifikasi pada sistem menggunakan metode *Naïve Bayes*, dimana data latih yang digunakan adalah data ulasan yang telah diberi label secara otomatis menggunakan kamus InSet yang telah melalui penyesuaian dengan data ulasan aplikasi Victoria Mobile Banking.

## II. LANDASAN KEPUSTAKAAN

### 2.1 Analisis Sentimen

Analisis sentimen adalah teknik *Natural Language Processing* (NLP) yang dapat digunakan untuk menentukan apakah suatu data bersifat netral, negatif, atau positif. Analisis sentimen juga dapat disebut sebagai *opinion mining* [12]. Analisis sentimen biasanya digunakan pada data berbentuk teks untuk membantu sebuah bisnis atau perusahaan dalam memonitor sentimen *customer feedback* terhadap brand itu sendiri ataupun suatu produk, dan juga membantu perusahaan untuk memahami kebutuhan pelanggan dan mengambil keputusan bisnis [5]. Dengan melakukan analisis sentimen, perusahaan memiliki keunggulan dalam memahami pengguna yang dapat berpengaruh pada penentuan teknik *marketing*, penentuan strategi bisnis, pengelolaan, pelayanan, dan lainnya [13], [14]. Wawasan atau *insight* yang didapat dari hasil analisis sentimen juga dapat dijadikan sebagai kunci strategi dalam membuat pengalaman pengguna (*customer experience*) yang baik, sehingga akan mendorong loyalitas

pelanggan serta meningkatkan strategi *marketing word-of-mouth* [15].

## 2.2 InSet (*Indonesia Sentiment Lexicon*)

*Indonesia Sentiment Lexicon*, atau yang dapat disingkat menjadi InSet, merupakan sebuah kamus sentimen lexicon berbahasa Indonesia yang memiliki 3.609 kata positif dan 6.609 kata negatif yang dapat diakses melalui GitHub. Kamus InSet dibuat berdasarkan data dari Twitter, dikarenakan pada saat penelitian dijalankan Twitter merupakan social media yang paling sering digunakan di Indonesia. Data Twitter yang digunakan untuk membuat kamus InSet sekitar 10.000 data yang dibagi menjadi data yang memiliki sentimen positif dan data yang memiliki sentimen negatif. Data tersebut kemudian melalui tahap preprocessing, lalu untuk pemilihan kandidat kata pada kamus, digunakan n-gram dengan  $n = \{1, 2, 3\}$ . Setelah itu tiap kata diberi pembobotan polaritas manual oleh kedua peneliti mulai dari -5 (sangat negatif) sampai dengan +5 (sangat positif). Dari penelitian tersebut, dapat diketahui bahwa penggunaan kamus InSet untuk penentuan sentimen berbahasa Indonesia lebih akurat dibandingkan menggunakan english lexicon yang ditranslasi, seperti Vania Lexicon, SentiWordNet, Liu Lex, dan AFINN, yaitu sebesar 65,78% [16]. Untuk memperoleh hasil yang lebih baik lagi dalam pengklasifikasian sentimen diperlukan penyesuaian konteks kamus InSet dengan *dataset* yang akan digunakan, yaitu data ulasan aplikasi Victoria Mobile Banking [17].

## 2.3 Naïve Bayes

*Naïve Bayes* adalah salah satu metode sederhana yang dapat melakukan klasifikasi dengan tingkat akurasi yang bagus untuk data yang berjumlah besar. Secara sederhana, metode *Naïve Bayes* mengasumsikan bahwa setiap atribut bersifat independen, sehingga kehadiran atribut tertentu tidak terkait dengan kehadiran atribut lainnya berdasarkan aturan *Bayes*. *Naïve Bayes* merupakan metode yang efektif dan efisien, sehingga sering kali digunakan untuk melakukan klasifikasi atau memberi prediksi secara *real-time* [18], [19]. Persamaan *Naïve Bayes* yang dapat dilihat dalam Persamaan (1) [20].

$$P(p|n) \propto P(p) \prod_{1 \leq k \leq n_d} P(t_k|p) \quad (1)$$

Keterangan:

- $n$  : Jumlah dokumen
- $p$  : Polaritas
- $n_d$  : Jumlah token dalam dokumen
- $t_k$  : Token ke- $k$  di dokumen
- $P(p)$  : Probabilitas *prior* dari dokumen yang memiliki polaritas  $p$
- $P(p|n)$  : Prediksi polaritas untuk pada sebuah dokumen
- $P(t_k|p)$  : Probabilitas munculnya suatu istilah ( $t_k$ ) dalam dokumen dengan polaritas  $p$

Persamaan untuk mendapatkan  $P(t_k|p)$  dapat dilihat pada Persamaan (2). Persamaan untuk mendapatkan  $P(p)$  dapat dilihat pada Persamaan (3).

$$P(t_k|p) = \frac{\text{count}(t_k|p) + 1}{\text{count}(t_p) + |V|} \quad (2)$$

Keterangan:

- $|V|$  : Jumlah istilah dalam total kosakata dokumen
- $\text{count}(t_k|p)$  : Jumlah kemunculan suatu istilah ( $t_k$ ) dalam dokumen yang memiliki polaritas  $p$
- $\text{count}(t_p)$  : Jumlah token yang ada dalam dokumen dengan polaritas  $p$

$$P(p) = \frac{\text{Number of documents of polarity } p}{\text{Total number of documents}} \quad (3)$$

## 2.4 Confusion Matrix

*Confusion matrix* merupakan sebuah metode yang dapat digunakan untuk melakukan evaluasi untuk menghitung tingkat akurasi dari sebuah proses klasifikasi [21]. *Confusion matrix* menggambarkan sebuah informasi mengenai klasifikasi aktual dan prediksi dari hasil proses sebuah sistem klasifikasi dengan bentuk dua dimensi yang merepresentasikan empat kombinasi dari nilai prediksi dan nilai aktual [22]. Pada Tabel 2.1 dapat dilihat bentuk *confusion matrix*.

**Tabel 2.1** *Confusion Matrix*

		Aktual	
		Negatif	Positif
Prediksi	Negatif	TN	FN
	Positif	FP	TP

Keterangan:

TP (*True Positive*) : Data aktual dan data prediksi bersifat positif

TN (*True Negative*) : Data aktual dan data prediksi bersifat negatif

FP (*False Positive*) : Data aktual bersifat negatif, namun pada data prediksi bersifat positif (*type I errors*)

FN (*False Negative*) : Data aktual bersifat positif, namun pada data prediksi bersifat negatif (*type II errors*)

Melalui tabel *confusion matrix*, dapat dilakukan perhitungan berdasarkan nilai TP, TN, FP, dan FN [23] untuk mendapat nilai *accuracy* (akurasi), *precision* (presisi), *recall* (penarikan), dan *f-measure* dengan persamaan sebagai berikut:

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

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

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

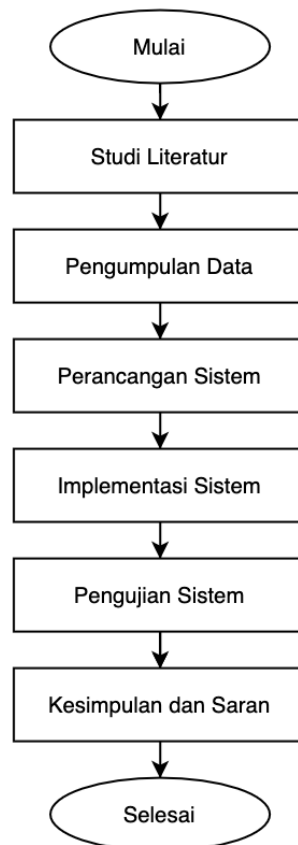
$$f - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (7)$$

Akurasi merepresentasikan jumlah data yang diklasifikasikan dengan benar dari data keseluruhan. Nilai akurasi dapat digunakan sebagai acuan dari performa sebuah algoritma atau sistem bila data yang digunakan dalam pengujian *confusion matrix* berimbang atau *balanced*. *Precision* (presisi) merepresentasikan jumlah prediksi data positif yang benar, atau dapat dikatakan sebagai tingkat akurasi dari prediksi data positif. *Recall* merepresentasikan jumlah data positif yang diklasifikasikan atau diprediksi dengan benar. *F-measure* merupakan *harmonic mean* atau rata-rata dari

nilai *recall* dan *precision*. Nilai *f-measure* hanya akan menjadi tinggi bila kedua nilai *recall* dan *precision* tinggi juga, sehingga nilai *f-measure* lebih baik digunakan sebagai acuan dari performa algoritma atau sistem bila data yang digunakan dalam pengujian *confusion matrix* tidak imbang atau *imbalanced* [24], [25].

### III. METODOLOGI

Metode penelitian pada penelitian ini terdiri dari beberapa tahapan. Diagram alir metodologi dalam penelitian ini dapat dilihat pada Gambar 3.1 berikut:



**Gambar 3.1** Diagram Alir Metodologi

#### 3.1 Studi Literatur

Pada tahap ini dilakukan studi literatur untuk mendapat gambaran mengenai penelitian analisis sentimen yang telah dilakukan oleh penelitian-penelitian sebelumnya dan juga dasar teori yang mendukung penelitian ini.

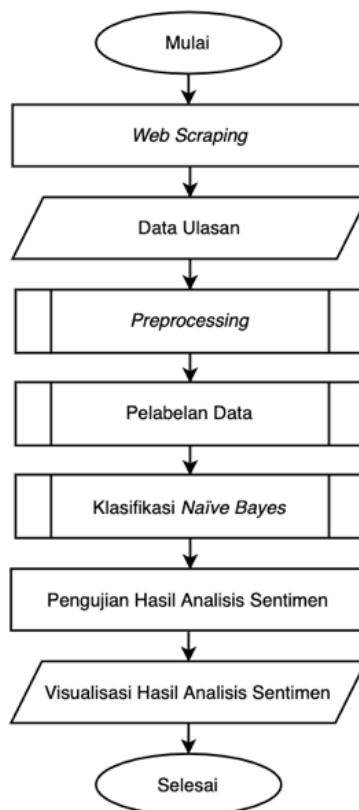
#### 3.2 Pengumpulan Data

Penelitian ini dilakukan dengan studi kasus aplikasi Victoria Mobile Banking, sehingga data yang digunakan adalah data ulasan aplikasi Victoria Mobile Banking yang ada pada Google Play Store Indonesia. Pengumpulan data pada tahap ini dilakukan dengan cara *web scraping* menggunakan *library google-play-scraper*. Hasil dari tahap pengumpulan data didapat 288 ulasan dari periode 2016 – 2022, dimana 34 ulasannya berbahasa Inggris (12% dari data ulasan keseluruhan).

#### 3.3 Perancangan Sistem

Tahap perancangan sistem menjelaskan mengenai gambaran cara kerja sistem yang akan dibuat dan juga perancangan sistem analisis sentimen. Perancangan yang dibuat akan berbentuk *workflow* diagram atau diagram alir yang menggambarkan tahapan proses yang dilalui untuk membuat sistem

analisis sentimen menggunakan metode *Naïve Bayes*.



**Gambar 3.2** Alur Sistem Analisis Sentimen

Sesuai dengan alur sistem analisis sentimen yang dapat dilihat pada Gambar 3.2, perancangan sistem yang dibuat pada tahap ini terdiri dari:

1. Perancangan *preprocessing* yang terdiri dari proses translasi dan formalisasi kata, *case folding*, *tokenizing*, *filtering*, dan *stemming* untuk mengolah dan menyesuaikan format data.
2. Perancangan pelabelan data dilakukan secara otomatis menggunakan kamus InSet (*Indonesia Sentiment Lexicon*). Kamus InSet telah melalui penyesuaian berupa penambahan kata dan perubahan bobot pada beberapa kata. Penambahan kata dilakukan berdasarkan kata yang memiliki *term frequency* lebih dari 1, namun belum tersedia di kamus. Penyesuaian bobot terhadap kata didasarkan pada dampak penggunaan kata tersebut terhadap hasil klasifikasi ulasan dan telah divalidasi oleh expert yang kompeten.
3. Perancangan klasifikasi *Naïve Bayes* terdiri dari pembagian data (empat skenario data uji yaitu 20%, 30%, 40%, dan 50%), *term weighting*, dan klasifikasi *Naïve Bayes*. Pembagian data dilakukan menggunakan parameter *random\_state* untuk menginisialisasi generator nomor acak internal, sehingga pembagian data secara *random* yang dihasilkan setiap kali sistem dijalankan tetap sama.
4. Perancangan pengujian *confusion matrix* untuk menampilkan tabel *confusion matrix*

### 3.4 Implementasi Sistem

Pada tahap implementasi sistem dilakukan pengimplementasian rancangan yang pada tahap sebelumnya telah dibuat. Proses implementasi sistem akan dilakukan menggunakan google colab dengan bahasa pemrograman *python*. Dalam proses implementasi sistem juga akan memanfaatkan beberapa *library* yang dibutuhkan dalam penelitian ini.

### 3.5 Pengujian Sistem

Pada tahap pengujian sistem akan dilakukan evaluasi atau pengujian menggunakan *confusion matrix*. Tahap pengujian ini dilakukan untuk mengetahui nilai *accuracy*, *precision*, *recall*, dan *f-measure* dari hasil klasifikasi *Naïve Bayes* yang didapat dari tahap sebelumnya.

### 3.6 Kesimpulan dan Saran

Pada tahap terakhir dalam penelitian ini, dilakukan penarikan kesimpulan dan pemberian saran

berdasarkan penelitian yang telah dilakukan.

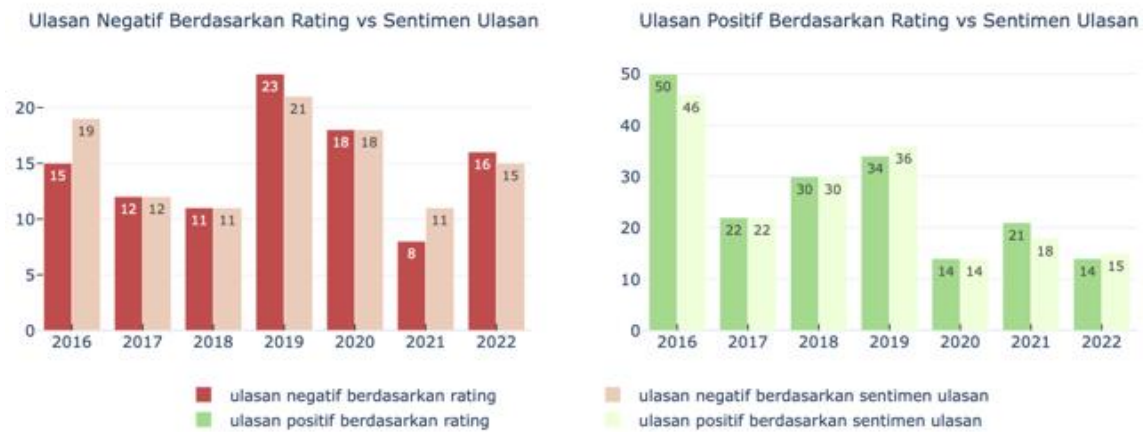
#### IV. PENGUJIAN DAN ANALISIS

Berdasarkan hasil pelabelan menggunakan kamus InSet pada studi kasus ulasan aplikasi Victoria Mobile Banking terdapat 173 ulasan negatif dan 115 ulasan positif, sedangkan berdasarkan hasil pelabelan menggunakan kamus InSet yang telah melalui penyesuaian dengan data ulasan didapat 181 ulasan positif dan 107 ulasan negatif dari 288 data ulasan aplikasi Victoria Mobile Banking.

##### 4.1 Hasil Visualisasi

Selain menghasilkan klasifikasi sentimen ulasan menjadi dua, yaitu positif atau negatif, sistem analisis sentimen juga menghasilkan visualisasi hasil analisis. Dari visualisasi hasil analisis yang berupa grafik ini diharapkan akan mempermudah pihak perusahaan dalam mengetahui dan memahami hasil dari analisis sentimen.

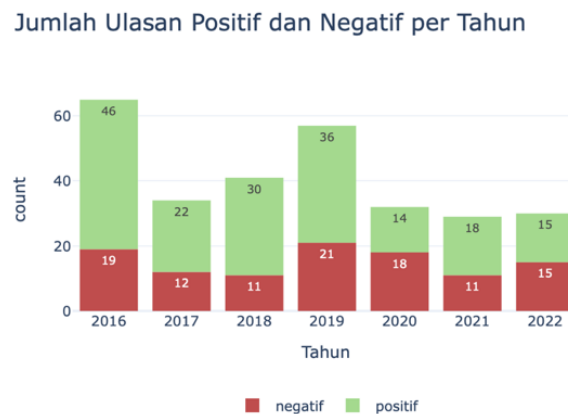
##### 4.1.1 Perbandingan Klasifikasi Berdasarkan Rating dan Sentimen Ulasan



**Gambar 4.1** Perbandingan Klasifikasi Berdasarkan *Rating* dan Sentimen Ulasan

Dalam sebuah ulasan terkadang terdapat ketidaksesuaian antara isi ulasan dengan pemberian bintang pada ulasan seperti yang telah ditulis pada latar belakang penelitian ini, yang berarti bahwa sentimen dari sebuah ulasan tidak dapat dilihat dari jumlah bintang atau *rating* yang diberikan oleh pengguna. Ketidaksesuaian tersebut terlihat pada kedua grafik pada Gambar 4.1. Ulasan yang diberi bintang tiga kebawah diumpamakan sebagai data negatif dan untuk ulasan yang diberi bintang empat dan lima akan diumpamakan sebagai data positif.

##### 4.1.2 Jumlah Ulasan Negatif dan Positif per Tahun

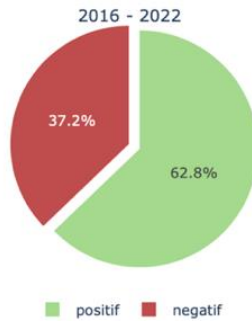


**Gambar 4.2** Jumlah Ulasan Negatif dan Positif per Tahun

Hasil visualisasi kedua pada sistem analisis sentimen penelitian ini adalah grafik yang menggambarkan jumlah data ulasan negatif dan positif per tahun yang dapat dilihat pada Gambar 4.2.

**4.1.3 Persentase Jumlah Ulasan Negatif dan Positif**

Persentase Jumlah Ulasan Positif & Negatif



**Gambar 4.3** Persentase Jumlah Ulasan Negatif dan Positif

Hasil visualisasi ketiga pada sistem analisis sentimen penelitian ini adalah grafik yang menggambarkan persentase jumlah data ulasan negatif dan positif pada aplikasi Victoria Mobile Banking yang dapat dilihat pada Gambar 4.3. Pada grafik tersebut dapat diketahui bahwa secara keseluruhan (dari tahun 2016-2022), pengguna aplikasi Victoria Mobile Banking lebih banyak memberikan ulasan positif dibandingkan dengan ulasan negatif, yaitu sebesar 64,2% ulasan positif.

**4.1.4 Word Cloud Ulasan Positif**



**Gambar 4.4** Word Cloud Positif Menggunakan InSet



**Gambar 4.5** Word Cloud Positif Menggunakan InSet yang Telah Melalui Penyesuaian

Hasil visualisasi keempat pada sistem analisis sentimen penelitian ini adalah *word cloud* ulasan positif. *Word cloud* merupakan sebuah gambar yang merepresentasikan kata-kata yang memiliki frekuensi muncul paling banyak pada sebuah data teks. Semakin besar sebuah kata pada *word cloud*, berarti bahwa semakin banyak kata tersebut ada pada sebuah data dan semakin penting juga kata tersebut. Gambar 4.4 merupakan gambar *word cloud* berdasarkan data ulasan yang diberi label menggunakan kamus InSet, sedangkan Gambar 4.5 merupakan gambar *word cloud* berdasarkan data



ulasan yang diberi label menggunakan kamus InSet yang sudah melalui penyesuaian dengan data ulasan.

Dapat dilihat pada Gambar 4.5, pada data ulasan positif beberapa kata yang paling besar dan yang paling terlihat pada *word cloud* adalah kata “bagus”, “mudah”, “bantu”, dan “transaksi”, yang berarti bahwa berdasarkan data ulasan positif, banyak pengguna aplikasi Victoria Mobile Banking yang berpendapat bahwa aplikasi tersebut sangatlah bagus, mudah digunakan, dan sangat membantu dalam melakukan transaksi

#### 4.1.5 Word Cloud Ulasan Negatif



Gambar 4.6 *Word Cloud* Negatif Menggunakan InSet



Gambar 4.7 *Word Cloud* Negatif Menggunakan InSet yang Telah Melalui Penyesuaian

Hasil visualisasi terakhir pada sistem analisis sentimen penelitian ini adalah *word cloud* ulasan negatif. Gambar 4.6 adalah gambar *word cloud* berdasarkan data ulasan yang diberi label menggunakan kamus InSet, sedangkan Gambar 4.7 adalah gambar *word cloud* berdasarkan data ulasan yang diberi label menggunakan kamus InSet yang sudah melalui penyesuaian dengan data ulasan.

Dapat dilihat pada Gambar 4.7 diatas, pada data ulasan negatif beberapa kata yang paling besar dan yang paling terlihat pada *word cloud* adalah kata “login”, “tolong”, “update”, “ribet”, dan “aktivasi”, yang berarti bahwa berdasarkan data ulasan negatif, beberapa pengguna aplikasi Victoria Mobile Banking merasa bahwa terdapat beberapa hal terkait proses *login*, *update*, ataupun aktivasi akun yang masih membingungkan dan memerlukan bantuan dari pihak bank.

## 4.2 Pengujian *Confusion Matrix* dan Analisis

### 4.2.1 Pengujian Data Hasil Klasifikasi Menggunakan Kamus InSet

Hasil pengujian *confusion matrix* dari data yang diklasifikasikan menggunakan kamus InSet dapat dilihat pada Tabel 4.1.

Tabel 4.1 *Confusion Matrix* (InSet)

Pembagian Data (Data Latih : Data Uji)	Prediksi	Aktual	
		Negatif	Positif
80% : 20%	Negatif	28	10

	<b>Positif</b>	1	19
<b>70% : 30%</b>	<b>Negatif</b>	47	13
	<b>Positif</b>	3	24
<b>60% : 40%</b>	<b>Negatif</b>	63	21
	<b>Positif</b>	2	30
<b>50% : 50%</b>	<b>Negatif</b>	78	32
	<b>Positif</b>	2	32

Berdasarkan Tabel 4.1 dapat diketahui bahwa kemampuan sistem klasifikasi dalam memprediksi data negatif lebih bagus dibandingkan dengan kemampuan dalam memprediksi data positif, dikarenakan pada data hasil pelabelan menggunakan kamus InSet lebih banyak data negatif dibanding data positif (dari 288 ulasan aplikasi Victoria Mobile Banking, 173 data berupa ulasan negatif dan 115 data berupa ulasan positif).

#### 4.2.2 Pengujian Data Hasil Klasifikasi Menggunakan Kamus InSet yang Telah Melalui Penyesuaian

Hasil pengujian *confusion matrix* dari data yang diklasifikasikan menggunakan kamus InSet yang telah disesuaikan dengan data ulasan dapat dilihat pada Tabel 4.2.

**Tabel 4.2** *Confusion Matrix* (InSet yang Telah Melalui Penyesuaian)

<b>Pembagian Data (Data Latih : Data Uji)</b>	<b>Prediksi</b>	<b>Aktual</b>	
		<b>Negatif</b>	<b>Positif</b>
<b>80% : 20%</b>	<b>Negatif</b>	16	0
	<b>Positif</b>	4	38
<b>70% : 30%</b>	<b>Negatif</b>	23	2
	<b>Positif</b>	12	50
<b>60% : 40%</b>	<b>Negatif</b>	28	4
	<b>Positif</b>	16	68
<b>50% : 50%</b>	<b>Negatif</b>	36	5
	<b>Positif</b>	15	88

Berdasarkan Tabel 4.2 dapat diketahui bahwa kemampuan sistem klasifikasi dalam memprediksi data positif lebih bagus dibandingkan dengan kemampuan dalam memprediksi data negatif, dikarenakan pada data hasil pelabelan menggunakan kamus InSet yang telah melalui penyesuaian lebih banyak data positif dibanding data negatif (dari 288 ulasan aplikasi Victoria Mobile Banking, 181 data berupa ulasan positif dan 107 data berupa ulasan negatif).

### 4.2.3 Analisis Hasil Pengujian

Tabel 4.3 Hasil Pengujian

	Pembagian Data (Data Latih : Data Uji)	<i>precision</i>	<i>recall</i>	<i>f-measure</i>	<i>accuracy</i>
		80% : 20%	<b>0.95</b>	<b>0.6551724</b>	<b>0.7755102</b>
Pelabelan Data Menggunakan InSet	70% : 30%	0.8888889	0.6486486	0.75	0.8160920
	60% : 40%	0.9375000	0.5882353	0.7228916	0.8017241
	50% : 50%	0.9411765	0.50	0.6530612	0.7638889
Pelabelan Data Menggunakan InSet yang Telah Melalui Penyesuaian	80% : 20%	<b>0.9047619</b>	<b>1</b>	<b>0.95</b>	<b>0.9310345</b>
	70% : 30%	0.8064516	0.9615385	0.8771930	0.8390805
	60% : 40%	0.8095238	0.9444444	0.8717949	0.8275862
	50% : 50%	0.8543689	0.9462366	0.8979592	0.8611111

Berdasarkan hasil pengujian pada Tabel 4.3 dapat diketahui bahwa perbandingan pembagian antara data latih dan data uji sebesar 80%:20% memiliki nilai *precision*, *recall*, *f-measure*, dan *accuracy* yang lebih baik, sehingga pada penelitian ini pembagian antara data latih dan data uji yang digunakan adalah 80% data latih dan 20% data uji.

Berdasarkan tabel *confusion matrix* dari data hasil klasifikasi yang tahap pelabelannya menggunakan kamus InSet dengan perbandingan 80% data latih dan 20% data uji yang dapat dilihat pada Tabel 4.1, dapat diketahui bahwa data negatif dan positif memiliki jumlah yang seimbang atau *balanced* (29 ulasan positif dan negatif), sehingga nilai akurasi (*accuracy*) dari hasil pengujian yang ada pada Tabel 4.3 dapat digunakan sebagai acuan dari performa sebuah algoritma atau sistem yaitu sebesar 0.8103448 atau 81%.

Berdasarkan tabel *confusion matrix* dari data hasil klasifikasi yang tahap pelabelannya menggunakan kamus InSet yang telah melalui penyesuaian dengan perbandingan 80% data latih dan 20% data uji yang dapat dilihat pada Tabel 4.2, dapat diketahui bahwa jumlah data negatif dan positif tidak seimbang atau *imbalanced*. Jumlah data positif hampir 2x lipat dari data negatif (38 data positif dan 20 data negatif), yang berarti bahwa nilai *accuracy* kurang bagus jika digunakan sebagai acuan dari performa algoritma atau sistem karena jika menggunakan data yang tidakimbang maka nilai *accuracy* dapat menginterpretasi hasil yang salah, sehingga dapat digunakan nilai *f-measure* yaitu sebesar 0.95 atau 95% berdasarkan hasil pengujian pada Tabel 4.3 sebagai acuan dari performa algoritma atau sistem, dikarenakan *f-measure* merupakan *harmonic mean* atau rata-rata dari *precision* (menggambarkan tingkat akurasi dari prediksi positif) dan *recall* (menggambarkan *true positive rate*).

## V. KESIMPULAN DAN SARAN

Dari penelitian ini dapat ditarik kesimpulan sebagai berikut:

1. Berdasarkan pengujian *confusion matrix* dengan perbandingan 80% data latih dan 20% data uji dapat diketahui bahwa performa dari sistem klasifikasi *Naïve Bayes* pada sistem analisis sentimen penelitian ini sebesar 0.95 atau 95%.
2. Berdasarkan hasil dari analisis sentimen terhadap ulasan aplikasi Victoria Mobile Banking, dapat diketahui bahwa sampai dengan tanggal 2 Januari 2023, mayoritas ulasan pada aplikasi Victoria Mobile Banking merupakan ulasan positif dimana banyak pengguna yang menganggap bahwa aplikasi Victoria Mobile Banking sangat bagus, mudah digunakan, dan sangat membantu dalam melakukan transaksi, walaupun beberapa pengguna juga merasa bahwa terdapat beberapa hal terkait proses *login*, *update*, ataupun aktivasi akun yang masih membingungkan dan memerlukan bantuan, namun secara keseluruhan pengguna merasa puas dengan aplikasi Victoria Mobile Banking.

3. Berdasarkan *feedback* dari pihak perusahaan terkait sistem analisis sentimen aplikasi Victoria Mobile, hasil visualisasi dari sistem analisis sentimen sangat mempermudah perusahaan dalam mengetahui unsur negatif atau keluhan yang dialami oleh pengguna aplikasi Victoria Mobile Banking, sehingga perusahaan dapat menyesuaikan aplikasi dengan kebutuhan pengguna dan mengembangkan aplikasinya untuk memberi layanan yang lebih berkualitas.

Saran yang dapat diberikan untuk penelitian selanjutnya:

1. Kamus InSet (*Indonesia Sentiment Lexicon*) dapat digunakan dalam melakukan pelabelan sentimen secara otomatis, namun perlu dilakukan penyesuaian beberapa nilai bobot atau penambahan kata pada kamus untuk menyesuaikan konteks kamus InSet dengan *dataset* yang digunakan.
2. Sistem dapat dikembangkan dengan menambahkan pengkategorian aspek pada ulasan kedalam kategori layanan, aplikasi, dan lainnya, sehingga dapat membantu perusahaan dalam mengetahui lebih *detail* mengenai aspek apa yang perlu diperbaiki dan dikembangkan lagi

#### DAFTAR PUSTAKA

- [1] D. M. E. D. M. Hussein, "A survey on sentiment analysis challenges," *Journal of King Saud University - Engineering Sciences*, vol. 30, no. 4, hlm. 330–338, Okt 2018, doi: 10.1016/j.jksues.2016.04.002.
- [2] M. Wankhade, A. C. S. Rao, dan C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artif Intell Rev*, vol. 55, no. 7, hlm. 5731–5780, Okt 2022, doi: 10.1007/s10462-022-10144-1.
- [3] X. Li, B. Zhang, Z. Zhang, dan K. Stefanidis, "A sentiment-statistical approach for identifying problematic mobile app updates based on user reviews," *Information (Switzerland)*, vol. 11, no. 3, Mar 2020, doi: 10.3390/info11030152.
- [4] O. Oyeboode, F. Alqahtani, dan R. Orji, "Using Machine Learning and Thematic Analysis Methods to Evaluate Mental Health Apps Based on User Reviews," *IEEE Access*, vol. 8, hlm. 111141–111158, 2020, doi: 10.1109/ACCESS.2020.3002176.
- [5] W. Medhat, A. Hassan, dan H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, hlm. 1093–1113, Des 2014, doi: 10.1016/J.ASEJ.2014.04.011.
- [6] R. N. Devita, H. W. Herwanto, dan A. P. Wibawa, "Perbandingan Kinerja Metode Naive Bayes dan K-Nearest Neighbor untuk Klasifikasi Artikel Berbahasa Indonesia," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 4, hlm. 427, Okt 2018, doi: 10.25126/jtiik.201854773.
- [7] R. R. Putra, M. E. Johan, dan E. R. Kaburuan, "A naïve bayes sentiment analysis for fintech mobile application user review in Indonesia," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 5. World Academy of Research in Science and Engineering, hlm. 1856–1860, 1 September 2019. doi: 10.30534/ijatcse/2019/07852019.
- [8] L. Zhang, K. Hua, H. Wang, G. Qian, dan L. Zheng, "Sentiment analysis on reviews of mobile users," dalam *Procedia Computer Science*, 2014, vol. 34. doi: 10.1016/j.procs.2014.07.013.
- [9] A. Yasar dan M. M. Saritas, "Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 7, no. 2, hlm. 88–91, Jan 2019, doi: 10.18201/ijisae.2019252786.
- [10] A. P. Wibawa, "Metode-metode Klasifikasi," *Prosiding Seminar Ilmu Komputer dan Teknologi Informasi*, vol. 3, no. 1, 2018.
- [11] R. L. Mustofa dan B. Prasetyo, "Sentiment analysis using lexicon-based method with naive bayes classifier algorithm on #newnormal hashtag in twitter," dalam *Journal of Physics: Conference Series*, Jun 2021, vol. 1918, no. 4. doi: 10.1088/1742-6596/1918/4/042155.
- [12] C. Puschmann dan A. Powell, "Turning Words Into Consumer Preferences: How Sentiment Analysis Is Framed in Research and the News Media," *Social Media and Society*, vol. 4, no. 3, Jul 2018, doi: 10.1177/2056305118797724.
- [13] K. Bannister, "Understanding Sentiment Analysis: What It Is & Why It's Used," *Brandwatch*, hlm. 1–13, 2015, [Daring]. Tersedia pada: <https://www.brandwatch.com/2015/01/understanding-sentiment-analysis/>
- [14] D. D. Albesta, M. L. Jonathan, M. Jawad, O. Hardiawan, dan D. Suhartono, "The impact of sentiment analysis from user on Facebook to enhanced the service quality," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 4, 2021, doi: 10.11591/ijece.v11i4.pp3424-3433.
- [15] C. Gallagher, E. Furey, dan K. Curran, "The application of sentiment analysis and text analytics to customer experience reviews to understand what customers are really saying," *International Journal of Data Warehousing and Mining*, vol. 15, no. 4. 2019. doi: 10.4018/IJDWM.2019100102.
- [16] F. Koto dan G. Y. Rahmaningtyas, "Inset lexicon: Evaluation of a word list for Indonesian sentiment analysis in microblogs," dalam *Proceedings of the 2017 International Conference on Asian Language Processing, IALP 2017*, Feb 2018, vol. 2018-January, hlm. 391–394. doi: 10.1109/IALP.2017.8300625.

- [17] R. Firdaus, I. A. #2, dan A. Herdiani, “Lexicon-Based Sentiment Analysis of Indonesian Language Student Feedback Evaluation,” *Indonesia Journal on Computing (Indo-JC)*, vol. 6, no. 1, hlm. 1–12, Mei 2021, doi: 10.34818/INDOJC.2021.6.1.408.
- [18] H. Zhang, “The optimality of Naive Bayes,” dalam *Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference, FLAIRS 2004*, 2004, vol. 2, hlm. 562–567.
- [19] S. Ray, “What is Naive Bayes algorithm?,” *Analytics Vidhya*, hlm. 1–8, 2017, [Daring]. Tersedia pada: <https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained/#:~:text=Naive Bayes Model,What is Naive Bayes algorithm%3F,presence of any other feature.>
- [20] G. Singh, B. Kumar, L. Gaur, dan A. Tyagi, “Comparison between Multinomial and Bernoulli Naive Bayes for Text Classification,” dalam *2019 International Conference on Automation, Computational and Technology Management, ICACTM 2019*, 2019. doi: 10.1109/ICACTM.2019.8776800.
- [21] A. Sharma, “Confusion Matrix in Machine Learning,” *Geeksforgeeks*, 2018, [Daring]. Tersedia pada: <https://www.geeksforgeeks.org/confusion-matrix-machine-learning/>
- [22] X. Deng, Q. Liu, Y. Deng, dan S. Mahadevan, “An improved method to construct basic probability assignment based on the confusion matrix for classification problem,” *Inf Sci (N Y)*, vol. 340–341, hlm. 250–261, Mei 2016, doi: 10.1016/j.ins.2016.01.033.
- [23] I. Düntsch dan G. Gediga, “Confusion Matrices and Rough Set Data Analysis,” dalam *Journal of Physics: Conference Series*, Mei 2019, vol. 1229, no. 1. doi: 10.1088/1742-6596/1229/1/012055.
- [24] T. Kanstrén, “A Look at Precision, Recall, and F1-Score,” *Towards Data Science*, 2020. <https://towardsdatascience.com/a-look-at-precision-recall-and-f1-score-36b5fd0dd3ec> (diakses 10 Februari 2023).
- [25] K. P. Shung, “Accuracy, Precision, Recall or F1?,” *Towards Data Science*, 2018. <https://towardsdatascience.com/accuracy-precision-recall-or-f1-331fb37c5cb9> (diakses 10 Februari 2023).