

## **Analisis Sentimen Aplikasi Dana Menggunakan Naïve Bayes Classifier dan Support Vector Machine**

**Gabriel Genaro Warow<sup>1</sup>, Henry Pandia<sup>2\*</sup>**

Sistem Informasi, Universitas Advent Indonesia, Bandung, Indonesia

\*Corresponding Author: pandiahenry@unai.edu

### **Abstract**

*The presence of financial technology (fintech) has introduced a new payment system that is practical, easily accessible, and efficient, such as the Dana application. However, to understand user acceptance of this application, sentiment analysis was required. This study aimed to analyze user opinions on the Dana application through their reviews and evaluate the performance of Naïve Bayes Classifier and Support Vector Machine algorithms in classifying sentiments. User review data was obtained through the Google Play Scraper API. From the sentiment analysis, it was concluded that the majority of users (84.7%) had a negative view of the Dana application, while only 8.8% provided positive sentiment, and 6.5% were neutral. Performance evaluation indicated that both Naïve Bayes Classifier and Support Vector Machine functioned effectively in classifying review data.*

**Keywords:** Naïve Bayes Classifier; Support Vector Machine; Dana Application

### **Abstrak**

Kehadiran teknologi finansial (*fintech*) telah menghadirkan sistem pembayaran baru yang praktis, mudah diakses, dan efisien seperti aplikasi Dana. Namun, untuk memahami penerimaan pengguna terhadap aplikasi tersebut, analisis sentimen diperlukan. Penelitian ini bertujuan untuk menganalisis opini pengguna terhadap aplikasi Dana melalui ulasan mereka, serta mengevaluasi kinerja algoritma *Naïve Bayes Classifier* dan *Support Vector Machine* dalam mengklasifikasikan sentimen. Data ulasan pengguna diperoleh melalui *API Google Play Scraper*. Dari analisis sentimen, disimpulkan bahwa sebagian besar pengguna (84,7%) memiliki pandangan negatif terhadap aplikasi *Dana*, sedangkan hanya 8,8% memberikan sentimen positif, dan 6,5% bersifat netral. Evaluasi kinerja menunjukkan bahwa baik *Naïve Bayes Classifier* maupun *Support Vector Machine* berfungsi efektif dalam mengklasifikasikan data ulasan.

**Kata kunci:** *Naïve Bayes Classifier; Support Vector Machine; Aplikasi Dana*

### **1. Pendahuluan**

Salah satu dampak positif dari kemajuan teknologi digital adalah hadirnya teknologi baru dalam bidang keuangan yang disebut dengan *Financial Technology* atau *Fintech*. *Fintech* merupakan penggabungan dari sistem keuangan dengan teknologi digital. *Fintech* memungkinkan adanya transaksi keuangan secara praktis, mudah diakses, nyaman, murah dan modern. Terdapat beberapa jenis *fintech* yang saat ini tersedia di Indonesia, seperti *online payment*, pinjaman *peer-to-peer*, *insuretech*, *aggregator*, *crowdfunding*, dan lainnya [1]. Hal ini membuat meningkatnya transaksi digital yang terjadi. Menurut Bank Indonesia, pada Agustus 2022 pembayaran digital mengalami peningkatan nilai transaksi uang elektronik yang tumbuh sebesar 43,24% secara tahunan [2] [3]. Aplikasi dompet digital Dana merupakan satu platform *online payment* yang digunakan secara luas di Indonesia. Banyaknya pengguna aplikasi dompet digital Dana membuat munculnya berbagai macam opini dalam bentuk ulasan, terdapat ulasan negatif, ulasan positif dan netral, dapat tercerminkan dengan *rating* yang diberikan pengguna. *Google Play Store* merupakan platform distribusi aplikasi resmi yang dikembangkan oleh *Google* yang memungkinkan pengguna dalam mencari dan mengunduh aplikasi untuk sistem *Android* [4]. *Google Play Store* menyediakan fitur bagi pengguna sebuah aplikasi untuk memberikan ulasan pada aplikasi dan layanan yang mereka gunakan [5]. Ulasan yang diberikan pengguna aplikasi dapat menunjukkan kepuasan, keluhan dan juga terkadang

merupakan ulasan yang tidak relevan [6]. Ulasan tersebut dapat menjadi umpan balik bagi penyedia layanan maupun pertimbangan bagi calon pengguna [7].

Penelitian ini menggunakan analisa sentimen untuk menganalisa ulasan diberikan pengguna aplikasi Dana di *Google Play Store*. Analisa sentimen sering pula disebut juga dengan *opinion mining*. Penelitian mengenai analisa sentimen aplikasi Dana telah dilakukan oleh beberapa peneliti, salah satunya dengan algoritma *Support Vector Machine* (SVM), dengan tingkat akurasi, presisi, serta *recall* yang tinggi [7]. Peneliti yang lain yang turut melakukan analisa sentimen terhadap aplikasi dompet digital, yang mana salah satunya adalah aplikasi Dana, peneliti tersebut menggunakan metode algoritma *Naive Bayes Classifier* (NBC) [8]. Oleh sebab itu maka peneliti tertarik melakukan analisa sentimen menggunakan algoritma NBC dan SVM. Metode NBC memiliki kelebihan ialah dapat melakukan pengklasifikasian data yang sesuai dengan kebutuhannya. Dengan sistem analisa sentimen ini diharapkan peneliti dapat melakukan analisa dengan baik [9]. Metode SVM juga dapat melakukan pengklasifikasian data yang menggunakan *machine learning (supervised learning)* yang memproyeksikan kelas berdasarkan pola *data training*, dengan fitur *hyperlane* sebagai pemisah antara kelas opini positif dan opini negatif [10]. Data ulasan aplikasi Dana akan diperoleh menggunakan teknik *Web-Scraping*. Pengambilan data menggunakan *API Google Play Scrapper*. *API* digunakan untuk mengekstraksi data informasi ulasan pengguna aplikasi di *Google Play Store* dengan lebih mudah tanpa ketergantungan eksternal [11].

Berdasarkan pembahasan diatas maka peneliti bermaksud untuk melakukan penelitian dengan judul Analisis Sentimen Ulasan Aplikasi Dana menggunakan metode *Naive Bayes Classifier* dan *Support Vector Machine*. Dengan adanya penelitian ini diharapkan bisa mengetahui opini positif atau negatif dari pengguna aplikasi Dana berdasarkan ulasan di *Google Play Store*, sehingga dapat menjadi bahan pertimbangan calon pengguna sebelum menggunakan aplikasi Dana, serta dapat memberikan bukti empiris sehingga dapat dijadikan bahan evaluasi bagi perusahaan penyedia aplikasi Dana untuk melakukan pengembangan berikutnya, sekaligus membandingkan ketiga metode tersebut untuk menemukan metode manakah yang memiliki kinerja yang paling baik, memiliki tingkat akurasi yang paling baik, kemudahan dalam proses pengolahan data penelitian, serta akan didapatkan informasi mengenai sentimen masyarakat.

Tujuan dari dilakukannya penelitian adalah mengetahui sentimen pengguna terhadap Dana yang dianalisis menggunakan algoritma NBC dan SVM, seberapa banyak sentimen positif dan negatifnya, turut juga mengetahui bagaimana kinerja metode algoritma NBC dan SVM yang diterapkan pada, serta membandingkan kinerja antara algoritma NBC dan SVM dalam menganalisis sentimen pengguna terhadap aplikasi Dana di *Google Play Store*. Hasil dari penelitian ini diharapkan dapat menyajikan informasi tentang analisis sentimen terhadap aplikasi Dana di *Google Play Store* menggunakan algoritma NBC dan SVM, mengetahui perbandingan kinerja dari kedua algoritma yang digunakan, serta dapat memberikan informasi kepada calon pengguna dan juga sebagai bahan evaluasi kepada perusahaan penyedia aplikasi Dana dalam melakukan pengembangan kualitas aplikasi Dana.

## 2. Tinjauan Pustaka

Terdapat beberapa penelitian terdahulu mengenai analisis sentimen terhadap aplikasi DANA telah dilakukan, namun terdapat perbedaan seperti jumlah dataset, waktu pengambilan data, serta uji algoritme yang digunakan.

Penelitian oleh Surohman dkk, dilakukan pengujian menggunakan algoritma *Naive Bayes Classifier* dan *K-Nearest Neighbors* (KNN), dengan hasil menunjukkan bahwa *Naive Bayes Classifier* lebih unggul dengan akurasi sebesar 84,76%, dibandingkan dengan KNN yang mencapai 82,92%. Meskipun demikian, penelitian ini belum mendalami faktor-faktor lain yang mungkin mempengaruhi sentimen pengguna, seperti pengalaman pengguna atau fitur-fitur aplikasi yang mungkin menjadi penentu [5].

Penelitian oleh Muhammad menggunakan metode *Support Vector Machine* (SVM) dan seleksi fitur chi square, menghasilkan akurasi sebesar 87,58% dan 89,41% untuk SVM dan SVM-*chi square*, masing-masing. Namun, belum ada penelitian yang mendalami area pemilihan fitur yang mungkin dapat meningkatkan performa model [7].

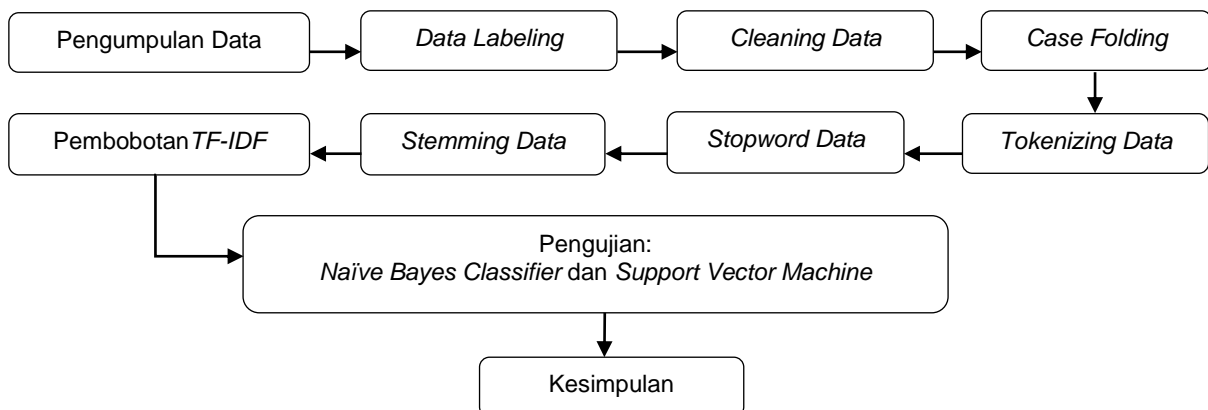
Penelitian oleh Harfian menggunakan *Naive Bayes Classifier* dalam mengklasifikasikan komentar pada halaman Instagram Dana, dengan akurasi mencapai 93,33% [6]. Namun, tidak

disebutkan apakah terdapat faktor-faktor tambahan yang bisa meningkatkan akurasi atau perbandingan dengan metode lain yang mungkin lebih efisien.

Meskipun telah ada banyak penelitian sebelumnya yang menginvestigasi analisis sentimen terhadap berbagai aplikasi di platform digital, namun terdapat sedikit penelitian yang secara khusus fokus pada aplikasi Dana di *Google Play Store*. Oleh karena itu, terdapat kekosongan pengetahuan dalam hal pemahaman tentang bagaimana pengguna merespons aplikasi Dana secara khusus. Selain itu, walaupun algoritma *Naive Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM) telah banyak digunakan dalam analisis sentimen, belum ada banyak perbandingan yang mendalam antara kinerja kedua metode tersebut dalam konteks aplikasi Dana. Oleh karena itu, terdapat kebutuhan akan penelitian yang mengisi celah pengetahuan ini dengan menganalisis sentimen pengguna terhadap aplikasi Dana menggunakan kedua metode algoritma tersebut, serta membandingkan kinerja mereka untuk memberikan wawasan yang lebih komprehensif. Penelitian ini memberikan manfaat langsung dengan memberikan pemahaman mendalam tentang respons pengguna terhadap aplikasi Dana di *Google Play Store*. Dengan membandingkan kinerja algoritma *Naive Bayes Classifier* (NBC) dan *Support Vector Machine* (SVM), penelitian ini memberikan panduan praktis bagi perusahaan pengembang dalam meningkatkan kualitas layanan mereka segera. Pengguna saat ini juga dapat memanfaatkan hasil penelitian ini sebagai referensi untuk memutuskan penggunaan aplikasi Dana berdasarkan evaluasi sentimen pengguna sebelumnya.

### 3. Metodologi

Penelitian ini dilakukan dengan langkah-langkah seperti yang ditunjukkan oleh Gambar 1. Proses penelitian dimulai dengan mengumpulkan data, *data labeling*, *cleaning data*, *case folding*, *tokenizing data*, *stopword data*, *stemming data*, pembobotan dan pengujian menggunakan algoritme yang dipilih. Penelitian kemudian diakhir dengan pengambilan kesimpulan.



Gambar 1 Langkah-langkah Penelitian

#### 3.1 Pengumpulan Data

Proses pengumpulan data dilakukan dengan cara mengumpulkan ulasan aplikasi Dana di *Google Play Store* menggunakan API *Google Play Scrapper*. Hasil scrapping kemudian ditempatkan dalam sebuah dataset dengan format file.csv. Proses Scrapping dilakukan pada tanggal 2 April 2024 dan diperoleh sebanyak 398 ulasan pengguna. Ulasan tersebut disampaikan pengguna dengan selang waktu dari tanggal 17 Januari 2024 sampai 2 April 2024.

Dari 398 ulasan diperoleh 277 ulasan memberikan rating 1, 60 orang memberikan rating 2, 26 orang memberikan rating 3, 6 orang memberikan rating 4 dan 29 orang memberikan rating 5. Dari 398 ulasan tersebut, aplikasi Dana mendapatkan rata-rata rating 1,62. Contoh data ulasan yang digunakan di penelitian ini dapat dilihat di Tabel 2.

#### 3.2 Data Labeling

*Data labeling* merupakan sebuah proses yang memberikan label positif dan negatif pada data ulasan yang ada. Peneliti akan melakukan pelabelan menggunakan teknik *crowdsourcing* dengan menggunakan beberapa orang membantu pelabelan untuk menghasilkan hasil yang maksimal [12]. Tabel 1 menyajikan contoh pemberian label pada data ulasan.

Tabel 1. Contoh Data Ulasan Hasil *Scraping*

No	Ulasan	Rating	Tanggal
1	Aplikasi bangke, gw top up dr bank BRI dgn nominal yg menurut gw ga sedikit yaitu 3.437.000 ga masuk, padahal transaksi berhasil, gw coba Lg 35rb langsung masuk.pas di tlp ga bisa ngasih solusi. Gw coba otak Atik restart aplikasinya, malah ga bisa kebuka skrg di hp gw. Gw coba buka di hp lain , bisa kebuka dan saldo gw udah seharian blm masuk juga. Gw saranin buat yg mau pake aplikasi ini mending mikir dua kali dulu deh, drpd nanti bernasib sama kaya gw.	1	2024-01-22
2	Setiap kali mau upgrade ke premium selalu susah,padahal semua data lengkap n jelas, ketika mau transfer uang kata ny suruh upgrade ke premium,sedangkan sudah upgrade,eh di persulit,ini gimana sich	1	2024-01-23
3	Semua langkah saran sudah dilakukan.. dari clear cache dan lain lain.. bahkan sampai coba download ulang.. Tapi tetap hasilnya sama.. aplikasi Mem-BeTe-kan.. tidak bisa kasih solusi yang jelas. Nomer bener.. OTP bener.. tapi susah masuk.. aneeeh.. haaddeeuuh.. pliss.. buatlah aplikasi yang membahagiakan bukan yang bikin BeTe.. 😞	1	2024-01-25
4	apk kirim atau terima uang, belanja vocer listrik, pulsa atau paket internet, bayar QRIS, TERBAIK utk saat ini. sistem pengamanan sdh modern krna menggunakan scan wajah. yg jelas gak pake ribet lah. 1 aja saran sy kalo bisa buat apk yg versi LITE nya jadi bagi pengguna yg memori internalnya kecil atau RAM nya kecil bisa juga menikmati apk ini. makasih dan sukses terus	5	2024-01-25
5	Pelayanan lewat DIANA lama dan suka dipersulit, suka error pula app nya kalo mau isi pulsa. Udah gitu kalo ada transaksi yg emang mencurigakan dan udh lapor jawabnya "kita tidak bisa memastikan untuk pengembalian uangnya" padahal bukti udah ada jelas kalo itu bukan kita yg ngelakuin, walaupun nominal kecil juga tetep aja bikin kesel.	1	2024-01-26
6	Update nya sering bug dan lemot. Sekarang lemot bgt pdhl pake wifi. Bisa2 nya yg muncul kosongan. Tanpa gambar dan tulisan. Kurang ringan dan praktis tidak seperti dahulu.	2	2024-01-27
7	Untuk upgrade premium dipersulit, sudah lapor customer service, data sudah di kirim lengkap dan jelas masih saja belum ada info. Bilangnya suruh tunggu 3-5 hari kerja tapi nyatanya sudah seminggu lebih belum ada kejelasan. Tolong profesional dong aplikasi sudah besar tp masalah seperti ini lambat penanganan nya!!!!	1	2024-01-27
8	Dana saya error terus ngga bisa lihat profil atau riwayat transfer, padahal pas awal awal di pakai baik baik aja lancar, tapi sekarang malah macet terus,udah di update sampai 2x tapi tetep nggk bisa, dibilang karena koneksi internet, padahal jaringan nya bagus 😞, mohon bantuannya 👍	3	2024-01-28
9	untuk versi saat ini sangat tidak nyaman banyak bug yang harus diperbaiki lagi. contohnya sering terjadi eror saat mau melakukan transaksi, yang lebih kesalnya lagi setiap setelah transfer mau lihat atau buka detail transaksi sangatlah susah dan terus menerus keterangan eror, memperlambat untuk membagikan bukti transfer.. tolong diperbaiki. terimakasih	2	2024-01-29
10	Sering lelet, apalagi pas banyak saldo. Leletnya minta ampun. Bisa seharian ga bs di pake, padahal top up dr m banking ke dana krn mau di pake transaksi, tp pas sdh di top up, aplikasinya malah ngeselin.	1	2024-01-29

Tabel 2 Proses Data Labeling

Ulasan	Label
Sangat mudah digunakan, transfer uang dengan cepat, dan fitur top up yang praktis. Saya suka Dana!	Positif
Saya sudah melakukan top up, tetapi uang tidak masuk ke saldo Dana saya. Sangat mengecewakan bgt!!!	Negatif

### 3.3 Data Preprocessing

*Data preprocessing* merupakan tahapan pemrosesan data sebelum dilakukan analisis. Proses ini terdiri dari proses *cleaning data*, *case folding*, *tokenizing data*, *stopword data* dan *stemming data* [13]. Berikut ini merupakan langkah-langkah *data preprocessing*:

### 1) *Cleaning Data*

*Cleaning data* adalah proses untuk membersihkan data dari duplikasi data, inkonsistensi, dan data yang salah. Proses ini mencakup proses menghapus tanda baca, angka, simbol, *link URL*, dan *username* di dalam teks [14] serta menghapus punctuasi dan nominal sehingga hanya menyisakan data yang berupa alfabet saja [12]. Dilakukan seperti contoh pada Tabel 3 berikut:

Tabel 3 Proses *Cleaning Data*

Teks Asli	Setelah <i>Cleaning</i>
Sangat mudah digunakan, transfer uang dengan cepat, dan fitur top up yang praktis. Saya suka Dana!	Sangat mudah digunakan transfer uang dengan cepat dan fitur top up yang praktis Saya suka Dana
Saya sudah melakukan top up, tetapi uang tidak masuk ke saldo Dana saya. Sangat mengecewakan bgt!!!	Saya sudah melakukan top up tetapi uang tidak masuk ke saldo Dana saya Sangat mengecewakan bgt

### 2) *Case Folding*

*Case folding* merupakan proses mengubah huruf kapital menjadi huruf kecil sehingga teks yang ada menjadi seragam [13]. Contoh proses *case folding* dapat dilihat di Tabel 4 di bawah ini:

Tabel 4 Proses *Case Folding*

Setelah <i>Cleaning</i>	Setelah <i>Case Folding</i>
Sangat mudah digunakan transfer uang dengan cepat dan fitur top up yang praktis Saya suka Dana	sangat mudah digunakan transfer uang dengan cepat dan fitur top up yang praktis saya suka dana
Saya sudah melakukan top up tetapi uang tidak masuk ke saldo Dana saya Sangat mengecewakan bgt	saya sudah melakukan top up tetapi uang tidak masuk ke saldo dana saya sangat mengecewakan bgt

### 3) *Tokenizing Data*

Proses *cleaning data* adalah proses memecah sebuah kalimat menjadi sekumpulan kata terpisah. Kalimat diuraikan menjadi kata-kata agar data dapat diekstraksi lebih lanjut dan digunakan pada penelitian ini [15]. Hasil proses *tokenizing data* dapat dilihat di Tabel 5.

Tabel 5 *Tokenizing Data*

Setelah <i>Case Folding</i>	Setelah <i>Tokenizing Data</i>
sangat mudah digunakan transfer uang dengan cepat dan fitur top up yang praktis saya suka dana	[sangat] [mudah] [digunakan] [transfer] [uang] [dengan] [cepat] [dan] [fitur] [top] [up] [yang] [praktis] [saya] [suka] [dana]
saya sudah melakukan top up tetapi uang tidak masuk ke saldo dana saya sangat mengecewakan bgt	[saya] [sudah] [melakukan] [top] [up] [tetapi] [uang] [tidak] [masuk] [ke] [saldo] [dana] [saya] [sangat] [mengecewakan] [bgt]

### 4) *Stopword Data*

*Stopword data* adalah proses menghapus imbuhan atau awalan dari setiap kata per kata yang telah dipisahkan dari proses sebelumnya, sehingga menghasilkan kata-kata dasar [14]. Hasil proses *stopword* ditunjukkan di Tabel 6.

Tabel 6 Proses *Stopword Data*

Setelah <i>Tokenizing Data</i>	Setelah <i>Stopword Data</i>
[sangat] [mudah] [digunakan] [transfer] [uang] [dengan] [cepat] [dan] [fitur] [top] [up] [yang] [praktis] [saya] [suka] [dana]	[sangat] [mudah] [guna] [transfer] [uang] [dengan] [cepat] [dan] [fitur] [top] [up] [yang] [praktis] [saya] [suka] [dana]
[saya] [sudah] [melakukan] [top] [up] [tetapi] [uang] [tidak] [masuk] [ke] [saldo] [dana] [saya] [sangat] [mengecewakan] [bgt]	[saya] [sudah] [laku] [top] [up] [tapi] [uang] [tidak] [masuk] [ke] [saldo] [dana] [saya] [sangat] [kecewa] [bgt]

### 5) *Stemming Data*

*Stemming Data* merupakan proses melakukan koreksi salah eja pada setiap kata di ulasan dengan format kata yang benar sehingga tidak ada lagi kesalahan makna [14]. Hasil proses stemming data dapat dilihat di Tabel 7.

Tabel 7 Proses *Stemming Data*

Setelah <i>Stopword Data</i>	Setelah <i>Stemming Data</i>
[sangat] [mudah] [guna] [transfer] [uang] [dengan] [cepat] [dan] [fitur] [top] [up] [yang] [praktis] [saya] [suka] [dana]	[sangat] [mudah] [guna] [transfer] [uang] [dengan] [cepat] [dan] [fitur] [top] [up] [yang] [praktis] [saya] [suka] [dana]
[saya] [sudah] [laku] [top] [up] [tapi] [uang] [tidak] [masuk] [ke] [saldo] [dana] [saya] [sangat] [kecewa] [bgt]	[saya] [sudah] [laku] [top] [up] [tapi] [uang] [tidak] [masuk] [ke] [saldo] [dana] [saya] [sangat] [kecewa] [bangett]

### 3.4 Pengujian

Peneliti menggunakan beberapa metode algoritma yaitu NBC dan SVM yang bertujuan peneliti akan sekaligus membandingkan ketiga metode tersebut untuk menemukan metode manakah yang memiliki kinerja yang paling baik, memiliki tingkat akurasi yang paling baik, kemudahan dalam proses pengolahan data penelitian, serta akan didapatkan informasi mengenai sentimen masyarakat. Setelah dilakukan pengumpulan data, berikutnya pembobotan kata *TF-IDF*, lalu analisis sentimen menggunakan metode NBC dan SVM.

### 3.5 *Naïve Bayes Classifier*

*Naïve Bayes Classifier* (NBC) biasanya digunakan pada penelitian penambahan teks dengan tujuan penyelesaian masalah pengklasifikasian objek. NBC memiliki tingkat akurasi pengolahan yang lebih tinggi dari algoritma yang lain dalam menentukan polaritas yang sebenarnya dari kalimat yang disajikan, bahkan dalam sekumpulan data yang tidak seimbang [14]. Metode ini juga memiliki keunggulan dari penggunaannya yang mudah, pemrosesan waktu yang singkat, dengan struktur klasifikasi yang sederhana, sehingga memiliki tingkat efektivitas dan efisiensi yang tinggi, walaupun saat digunakan pada *dataset* dengan jumlah besar [8].

Didukung oleh penelitian yang dilakukan oleh Basryah yang melakukan analisis sentimen pada aplikasi dompet digital di *Google Play Store*, dimana salah satunya ialah aplikasi Dana dengan menggunakan algoritma NBC. Dengan metode tersebut pada penelitian ini, didapatkan bahwa implementasi algoritma NBC memiliki keakurasian diatas 90% [8]. Namun NBC memiliki beberapa kekurangan, metode ini tidak dapat mengukur satu probabilitas saja, yang menyebabkan tingkat akurasi bukan ukuran yang benar [15]. Metode ini juga rentan pada banyak fitur yang dapat mempengaruhi tingkat akurasi yang menjadi rendah [14]. Secara lebih sederhana, Teorema Bayes merupakan cara untuk menghitung probabilitas satu peristiwa jika kita sudah mengetahui probabilitas peristiwa lainnya.

$$P(A | B) = \frac{P(B | A) \times P(A)}{P(B)} \dots\dots\dots (1)$$

dimana:

- 1)  $P(A | B)$ : Probabilitas dari kejadian A terjadi jika kejadian B telah terjadi. Ini merupakan probabilitas bersyarat dari A terhadap B.
- 2)  $P(B | A)$ : Probabilitas dari kejadian B terjadi jika kejadian A telah terjadi. Ini juga merupakan probabilitas bersyarat dari B terhadap A.
- 3)  $P(A)$ : Probabilitas dari kejadian A terjadi secara independen tanpa mempertimbangkan kejadian B.
- 4)  $P(B)$ : Probabilitas dari kejadian A terjadi secara independen tanpa mempertimbangkan kejadian B.

Ketika digunakan untuk melakukan klasifikasi dalam sebuah analisa sentimen, langkah-langkah kerja algoritme NBC adalah sebagai berikut:

#### 1) Persiapan Data

Hal ini meliputi pengumpulan data, *data preprocessing*, dan *split data* menjadi *data training* dan *data testing*. *Data training* digunakan untuk melatih model dan *data testing* digunakan untuk menguji model.

#### 2) Ekstraksi Fitur

Setelah data siap, fitur atau atribut yang akan digunakan dalam pembelajaran mesin diekstraksi dari data. Dalam analisis sentimen menggunakan pembobotan kata *TF-IDF*.

### 3) Pembuatan Model

Model *Naive Bayes* dibuat dengan menggunakan *library scikit-learn* di *Python*. Pembuatan model melibatkan beberapa tahap, yaitu:

- a. Inisialisasi model dengan menggunakan *Multinomial Naive Bayes Classifier*
- b. Melatih model dengan menggunakan *data training* dan fitur yang telah diekstraksi
- c. Menguji model dengan menggunakan *data testing*

### 4) Evaluasi Model

Setelah model diuji, langkah selanjutnya adalah mengevaluasi kinerja model. Beberapa metrik evaluasi yang dapat digunakan dalam analisis sentimen, antara lain akurasi, presisi, recall, dan F1-score.

## 3.6 Support Vector Machine

*Support Vector Machine* (SVM) merupakan metode pengklasifikasian dengan dasar teoritis yang kuat serta memiliki tingkat keberhasilan empiris yang sangat baik. SVM merupakan pendekatan studi yang memanfaatkan fungsi linear dalam ruang dimensi tinggi dan dipelajari dengan algoritma berdasarkan teori optimisasi, dengan penerapan *learning bias* yang berasal dari teori statistik. Tujuan dari metode ini adalah untuk membangun *Optimal Separating Hyperlane*, dengan membuat fungsi pemisahan yang dapat digunakan untuk klasifikasi [15]. SVM memiliki keunggulan tingkat kinerja yang baik untuk diimplementasikan pada *bioinformatics*, pengenalan, tulisan tangan, klasifikasi dan lain sebagainya [14].

Algoritme SVM juga memiliki keunggulan yaitu dapat mengklasifikasikan suatu *pattern*, yang tidak termasuk data yang digunakan, serta metode ini mudah untuk diimplementasikan. Namun terdapat kekurangan yaitu sulit diimplementasikan pada fenomena dengan jumlah sampel yang besar, dan jenis kernel SVM berpengaruh terhadap tingkat akurasi sistem [14]. Penelitian lain telah dilakukan menggunakan algoritma SVM untuk menganalisis sentimen aplikasi pembelajaran *online* yang terdapat di *Google Play Store*, setelah dilakukan proses analisis terhadap 1500 sampel, metode ini dapat bekerja dengan baik, menghasilkan tingkat akurasi yang cukup tinggi dalam menentukan aplikasi yang memiliki sentimen positif [10]. Berikut adalah langkah-langkah kerja algoritma SVM saat melakukan analisis sentimen menggunakan metode di *Python*:

#### 1) Persiapkan dataset

#### 2) Melakukan *data preprocessing*

#### 3) *Bag-of-words representation*

Mengubah dataset ke dalam bentuk representasi *bag-of-words*. Representasi ini akan digunakan sebagai fitur untuk SVM.

#### 4) *Split* dataset menjadi data latih dan data uji

Dataset yang telah diolah kemudian perlu dibagi menjadi data latih dan data uji.

#### 5) *Training model SVM*

Setelah dataset dibagi, lakukan training model SVM dengan menggunakan data latih yang sudah dibuat. Pilih kernel dan hyperparameter yang tepat untuk meningkatkan performa model.

#### 6) Evaluasi model

Setelah model dilatih, evaluasi performanya dengan menggunakan data uji. Hitung akurasi, presisi, *recall*, dan *f1-score*. Berikut penjelasan tentang teknik-teknik dalam evaluasi:

##### a. Akurasi

Akurasi adalah salah satu aspek penting yang dinilai dalam evaluasi penelitian. Akurasi mengacu pada sejauh mana data dan hasil penelitian sesuai dengan realitas atau fenomena yang diteliti. Semakin tinggi tingkat akurasi maka semakin dapat dipercaya hasil penelitian tersebut.

##### b. Presisi

Presisi adalah salah satu aspek penting dalam evaluasi penelitian yang berkaitan dengan tingkat ketepatan atau keakuratan pengukuran atau pengamatan yang dilakukan dalam penelitian. Semakin tinggi tingkat presisi maka semakin akurat pula data dan hasil penelitian yang dihasilkan.

##### c. *Recall*

*Recall* penting dalam evaluasi penelitian yang berkaitan dengan kemampuan system atau model untuk mengenali dan mengingat Kembali data atau informasi yang relevan dengan masalah yang diteliti. *Recall* sering digunakan dalam penelitian yang berkaitan dengan pengolahan Bahasa alami, klasifikasi data, atau system rekomendasi.

d. *F1-score*

*F1-score* adalah salah satu metrik dalam evaluasi penelitian yang digunakan untuk mengukur keseluruhan kinerja system atau model dalam memprediksi atau mengklasifikasikan data. *F1-score* adalah rata-rata harmonis antara presisi dan recall. Dalam melakukan evaluasi penelitian *f1-score* digunakan untuk mengukur keseluruhan kinerja system atau model dalam memprediksi atau mengklasifikasikan data. Dalam kesimpulannya *f1-score* memberikan gambaran yang jelas tentang kinerja system atau model dalam memprediksi atau mengklasifikasikan data. Oleh karena itu dalam melakukan evaluasi penelitian, penilaian terhadap *F1-score* harus diperhatikan dengan cermat dan hati-hati. Setiap baris dari matrix, merepresentasikan kelas aktual dari data, dan setiap kolom merepresentasikan kelas prediksi dari data (atau sebaliknya). *Matrix* tersebut dijelaskan pada Tabel 8 berikut:

Tabel 8 *F1-Score*

	Predicted Negatif	Predicted Positif
Actual Negatif	True Negatif (TN)	False Positif (FP)
False (Negatif)	False Negatif (FN)	True Positif (TP)

Keterangan

TP = Jumlah data yang sebenarnya positif dan diprediksi positif

TN = Jumlah data yang sebenarnya negatif dan diprediksi negatif

FP = Jumlah data yang sebenarnya negatif tetapi diprediksi positif

FN = jumlah data yang sebenarnya positif tetapi diprediksi negatif.

Dengan menggunakan keempat data tersebut, dapat dihasilkan data lain yang sangat penting untuk menilai kinerja suatu model. Berikut rumus dari akurasi, presisi, *recall*, dan *F1-score*

Akurasi = Total keseluruhan seberapa sering model benar mengklasifikasikan.

Formula akurasi dapat ditulis sebagai berikut:

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

Presisi = Ketika model memprediksi positif, seberapa sering prediksi itu benar.

Formula presisi dapat ditulis sebagai berikut:

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

Recall = Ketika kelas aktualnya positif, seberapa sering model memprediksi positif.

Formula recall dapat ditulis sebagai berikut:

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

*F1-score* = Merupakan rata-rata harmonik dari presisi dan recall. Formula *f1-score* dapat ditulis sebagai berikut:

$$2 * (\text{Presisi} * \text{Recall} / \text{Presisi} + \text{Recall})$$

### 3.7 Pembobotan TF-IDF

Metode ini melibatkan pembobotan kata untuk meningkatkan analisis sentimen pada text mining, dengan *TF* (*Term Frequency*) memperhitungkan proporsi kemunculan kata dalam teks dan *IDF* (*Inverse Document Frequency*) memantau kemunculan token dalam koleksi teks. [14]. Berikut rumus dari TF-IDF:

$TF = (\text{jumlah kemunculan kata dalam dokumen}) / (\text{jumlah kata dalam dokumen})$

$IDF = \log (\text{jumlah seluruh dokumen dalam koleksi} / \text{jumlah dokumen yang mengundang kata tersebut})$

$TF-IDF = TF * IDF$

Hasil perhitungan TF-IDF akan menentukan apakah sebuah kata merupakan fitur yang penting dalam sebuah dokumen. Kata dengan nilai TF-IDF tinggi akan dianggap sebagai fitur



penting dari sebuah dokumen sebaliknya data dengan TF-IDF rendah akan dianggap sebagai kata yang kurang penting.

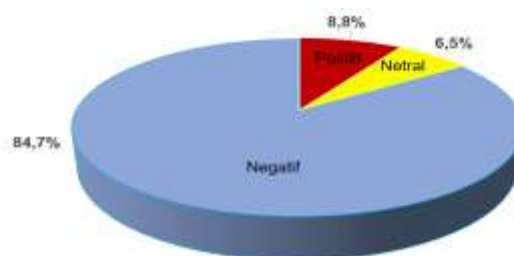
**4. Hasil dan Pembahasan**

Pada bagian ini berisikan proses pengujian, hasil serta pembahasan analisis sentimen aplikasi Dana menggunakan algoritma NBC dan SVM, menggunakan ulasan pengguna terbaru Dana di *Google Play Store* yang dikumpulkan menggunakan *Google Play Scraper*. Selanjutnya dilakukan *data preprocessing* diantaranya *data labeling*, *data cleaning*, *case folding*, *stopword removal*, *data tokenizing* dan *data stemming*. Tabel 9 berikut menampilkan 10 sampel *data preprocessing*:

Tabel 9 *Data Preprocessing*

	<i>content</i>	<i>score</i>	<i>Label</i>	<i>text_clean</i>	<i>text_StopWord</i>	<i>text_tokens</i>	<i>text_stemindo</i>
1	Setelah di updet bukannya makin lancar malah s...	2	Negatif	setelah di udet bukannya makin lancar malah s...	updet lancar aplikasi eror	[updet, lancar, aplikasi, eror]	updet lancar aplikasi eror
2	Sy barusan,isi,saldo k akun dana sy...tp pas s...	1	Negatif	sy barusanisaldo k akun dana sytp pas sy cek...	sy barusanisaldo k akun dana sytp pas sy cek...	[sy, barusanisald o, k, akun, dana, sytp, pas...]	sy barusanisaldo k akun dana sytp pas sy cek...
3	Mas tiap minggu update mulu. Ngabisin kuota aja	1	Negatif	mas tiap minggu update mulu ngabisin kuota aja	mas minggu update mulu ngabisin kuota aja	[mas, minggu, update, mulu, ngabisin, kuota, aja]	mas minggu update mulu ngabisin kuota aja
4	baguss	5	Positif	baguss	baguss	[baguss]	baguss
5	Kurang bagus	5	Positif	kurang bagus	bagus	[bagus]	bagus
6	Kenapa lah pas mau beli kuota di apk notifnya ...	2	Negatif	kenapa lah pas mau beli kuota di apk notifnya ...	pas beli kuota apk notifnya saldo dana mencuku...	[pas, beli, kuota, apk, notifnya, saldo, dana,...]	pas beli kuota apk notifnya saldo dana cukup i...
7	Layanan cs nya buruk sekali, chat bot tidak bi...	1	Negatif	layanan cs nya buruk sekali chat bot tidak bis...	layanan cs nya buruk chat bot mengerti hubungk...	[layanan, cs, nya, buruk, chat, bot, mengerti,...]	layan cs nya buruk chat bot erti hubung cs dan...
8	oke	5	Positif	oke	oke	[oke]	oke
9	Apk itl	1	Negatif	apk itl	apk itl	[apk, itl]	apk itl
10	bagus	5	Positif	bagus	bagus	[bagus]	bagus

Pada penelitian ini, polaritas sentimen dikelompokkan menjadi dua kategori utama sentimen, yakni positif serta negatif. Ulasan dengan sentimen positif, menunjukkan kegembiraan, kepuasan, atau kesan yang baik dari pengguna terhadap layanan aplikasi Dana. Sebaliknya, sentimen negatif menunjukkan kekecewaan, kemarahan, atau perasaan tidak puas dari pengguna terhadap layanan aplikasi Dana. Sedangkan tinjauan dengan polaritas netral tidak menunjukkan adanya sentimen positif atau negatif yang signifikan.



Gambar 2 Persentase Distribusi Sentimen

Gambar diatas menunjukkan persentase distribusi sentimen pengguna aplikasi Dana. Dimana sentimen positif disimbolkan angka “1”, sentimen negatif disimbolkan angka “-1” dan sentimen yang bersifat netral disimbolkan angka “0”. Dapat dilihat pada gambar diatas bahwa sentimen positif mendominasi dengan 67,2%. Lalu sentimen negatif dengan 28,7% dan sentimen yang bersifat netral dengan 4,7%. Mayoritas pengguna Dana memiliki sentimen positif

terhadap layanan Dana. Meskipun mayoritas pengguna aplikasi Dana menyatakan pandangan positif terhadap layanan tersebut, pengalaman pengguna tidak selalu berjalan mulus. Tidak sedikit pengguna menghadapi kesulitan karena sistem sering mengalami *error* atau *bug*, yang pada akhirnya menyebabkan ketidakpuasan. Namun, mayoritas ulasan menunjukkan bahwa ada pengguna Dana yang merasa puas dengan layanan tersebut.

Hasil pengujian sentimen dapat juga ditampilkan menggunakan *word cloud* yang dapat menampilkan kata yang paling sering muncul. Kata dengan ukuran yang besar menunjukkan kata tersebut relatif lebih sering muncul dibanding dengan kata yang lain. Word cloud untuk sentimen positif dan negatif pada penelitian ini ditunjukkan oleh Gambar 3 di bawah ini.



Gambar 3 Word cloud sentimen positif dan negatif

#### *Naïve Bayes Classifier*

Data yang digunakan pada penelitian ini diperoleh dengan melakukan *scrapping* pada ulasan pengguna aplikasi Dana. Data disimpan dalam bentuk file CSV. Data kemudian mengalami rangkaian proses pengolahan seperti yang sudah dijelaskan di atas. Data kemudian dipisahkan menjadi dua bagian, dimana 70 % data digunakan sebagai data train atau sebanyak 662 data. Sedangkan 30 % atau 285 data digunakan sebagai data test. Selanjutnya dilakukan proses analisa sentimen menggunakan algoritme NBC. Hasil analisa sentimen menggunakan algoritme NBC ditunjukkan di Gambar 4.

Accuracy: 0.9177631578947368				
Presisi: 0.9276440712924846				
Recall: 0.9177631578947368				
F1-score: 0.9149048579123875				
Confusion Matrix:				
[[ 81 10 15]				
[ 0 109 0]				
[ 0 0 89]]				
	precision	recall	f1-score	support
Negatif	1.00	0.76	0.87	106
Netral	0.92	1.00	0.96	109
Positif	0.86	1.00	0.92	89
accuracy			0.92	304
macro avg	0.92	0.92	0.91	304
weighted avg	0.93	0.92	0.91	304

Gambar 4 Hasil pengujian algoritme NBC

Evaluasi algoritma NBC menunjukkan hasil yang memuaskan dalam mengklasifikasikan sentimen pada *dataset*. Hasil pengujian menunjukkan algoritme NBC memberikan tingkat akurasi 91.78%, yang berarti hasil klasifikasi dapat diterima. Nilai presisi sebesar 92,78% mengindikasikan bahwa sekitar 92,78% dari prediksi sentimen positif yang dibuat oleh model adalah prediksi yang tepat. Nilai *recall* sebesar 91.78% menunjukkan kemampuan model dalam mengidentifikasi sebagian besar contoh sentimen positif yang sebenarnya dalam *dataset*. Nilai *F1-score*, yang merupakan rata-rata harmonik antara presisi dan *recall*, mencapai 91,49%, menandakan keseimbangan yang baik antara presisi dan *recall*.

*Confusion matrix* memberikan informasi tambahan tentang kinerja model pada setiap kelompok sentimen. Dalam kasus ini, terlihat bahwa NBC memiliki tingkat kesalahan yang lebih tinggi dalam memprediksi sentimen negatif, dengan hanya 76% prediksi yang benar (*recall*). Namun, model ini memiliki kinerja yang sangat baik dalam memprediksi sentimen positif dan netral dengan nilai *recall* 100%.

#### Support Vector Machine

Pada metode ini, pengambilan data juga dilakukan menggunakan *google-play-scraper*. Dimana selanjutnya juga dilakukan *data preprocessing* dan juga pembobotan TD-IDF menggunakan *TfidfVectorizer* serta dilakukan *split data* dengan perbandingan 70% *data\_train* dan 30% *data\_test*. Berikut ini hasil yang didapat setelah melakukan penelitian dengan algoritma klasifikasi algoritma *Support Vector Machine*:

```

Accuracy Score: 0.9177631578947368
Presisi: 0.9276440712924846
Recall: 0.9177631578947368
F1-score: 0.9149048579123875
Confusion Matrix:
[[ 81 10 15]
 [ 0 109 0]
 [ 0 0 89]]
=====
          precision  recall  f1-score  support
Negatif          1.00    1.00    1.00    106
Netral           1.00    1.00    1.00    109
Positif          1.00    1.00    1.00     89

 accuracy              1.00    304
 macro avg           1.00    1.00    1.00    304

```

Gambar 5 Hasil pengujian algoritme SVM

Evaluasi di atas menunjukkan algoritme SVM bekerja yang sangat baik dalam melakukan klasifikasi pada dataset yang diberikan. SVM memperoleh nilai presisi pada level 0,928, yang mengindikasikan bahwa sekitar 92.8% dari semua prediksi positif yang dibuat oleh model SVM adalah benar. *Recall* mengukur proporsi dari contoh positif yang telah diidentifikasi dengan benar dari total jumlah contoh yang sebenarnya positif. Pada evaluasi ini, recall mencapai nilai sekitar 0.918, yang berarti sekitar 91.8% dari semua contoh positif telah berhasil diidentifikasi oleh model.

F1-score adalah rata-rata harmonik dari presisi dan recall. SVM mempunyai nilai F1-score 0.915 yang menunjukkan keseimbangan antara presisi dan recall. Ini berarti model memiliki kinerja yang baik dalam mengklasifikasikan kelas positif dan negatif. *Confusion matrix* memberikan gambaran lebih rinci tentang kinerja model dalam mengklasifikasikan setiap kelas. Nilai *confusion matrix* menunjukkan sebagian besar prediksi benar, dimana sebagian besar elemen di diagonal utama memiliki nilai yang tinggi.

Selain itu, laporan klasifikasi menunjukkan bahwa model memiliki akurasi yang sempurna (1.00), yang mengindikasikan bahwa model SVM berhasil mengklasifikasikan keseluruhan data uji dengan benar. Secara keseluruhan, evaluasi tersebut menunjukkan bahwa model SVM sangat baik dalam melakukan klasifikasi pada dataset yang diberikan, dengan nilai presisi, recall, dan F1-score yang tinggi, serta akurasi yang sempurna.

Berdasarkan hasil evaluasi kinerja, ditemukan adanya perbedaan kinerja dan karakteristik algoritma kedua algoritma. Pertama, dari segi akurasi, SVM memiliki performa yang lebih baik dengan akurasi sempurna (1.00), sementara NBC mencapai akurasi sebesar 91.78%. Hal ini menunjukkan bahwa SVM mampu mengklasifikasikan semua data uji dengan benar, sedangkan NBC memiliki sedikit kesalahan dalam melakukan klasifikasi. Kedua, dari segi presisi, SVM dan NBC keduanya memiliki nilai yang tinggi, namun SVM sedikit lebih unggul dengan presisi sekitar 92.8%, sedangkan NBC memiliki presisi sekitar 92.76%. Ini menunjukkan bahwa sebagian besar prediksi positif yang dibuat oleh kedua model adalah benar.

Ketiga, dari nilai recall, SVM dan NBC memiliki kinerja yang baik, namun NBC memiliki sedikit keunggulan dengan nilai *recall* sekitar 91.78% dibandingkan dengan recall SVM sekitar 91.8%. Ini menunjukkan bahwa NBC sedikit lebih baik dalam mendeteksi sentimen positif dari total jumlah contoh yang sebenarnya positif. Keempat, *F1-score* yang merupakan rata-rata harmonik dari presisi dan recall menunjukkan keseimbangan antara kedua algoritme. Meskipun perbedaan nilainya sangat kecil, NBC memiliki *F1-score* sekitar 91.49%, sedangkan SVM memiliki *F1-score* sekitar 0.915.

Kelima, melalui confusion matrix, terlihat bahwa kedua model memiliki tingkat kesalahan yang rendah, namun NBC memiliki sedikit kesulitan dalam memprediksi sentimen negatif, sementara SVM berhasil mengklasifikasikan semua kelas dengan benar. Secara keseluruhan, SVM dan NBC menunjukkan kinerja yang sangat baik dalam melakukan klasifikasi pada dataset yang diberikan.

Namun, hasil penelitian ini tidak selaras dengan penelitian yang dilakukan oleh Nurian dan Sari, yang menyatakan bahwa aplikasi DANA mendapatkan respon positif dari pengguna di *Google Play Store* [16], serta tidak sejalan dengan penelitian yang dilakukan oleh Basryah [8], yang juga menyatakan bahwa aplikasi DANA mendapatkan respon yang positif dari pengguna. Perbedaan hasil tersebut mungkin disebabkan oleh waktu pengambilan data yang berbeda. Pengambilan data penelitian ini dilakukan pada periode 17 Januari 2024 hingga 2 April 2024 saja, sehingga mungkin saja pada periode tersebut, pengguna mengalami kendala dalam menggunakan aplikasi DANA, seperti yang tercermin dalam ulasan di *Google Play Store*. Selain itu, perbedaan hasil antara penelitian ini dengan penelitian sebelumnya dapat juga dipengaruhi oleh faktor-faktor lain, seperti perubahan dalam fitur atau fungsionalitas aplikasi DANA dari waktu ke waktu, perubahan preferensi atau ekspektasi pengguna terhadap aplikasi, serta perubahan dalam kondisi pasar atau persaingan dengan aplikasi sejenis. Penting untuk diingat bahwa hasil analisis sentimen tidak hanya bergantung pada algoritma yang digunakan, tetapi juga pada kualitas dan representativitas dataset yang digunakan dalam penelitian. Penelitian ini menggunakan dataset ulasan pengguna dari *Google Play Store*, namun tidak selalu mencakup semua pengguna atau semua aspek penggunaan aplikasi DANA. Oleh karena itu, hasil analisis sentimen mungkin tidak sepenuhnya mencerminkan pendapat seluruh pengguna atau pengalaman pengguna secara keseluruhan.

## 5. Simpulan

Berdasarkan analisis data sentimen terhadap layanan Dana, dapat disimpulkan bahwa mayoritas pengguna mengekspresikan sentimen negatif, yang mencapai 84,7%. Hal ini menandakan adanya ketidakpuasan yang cukup signifikan terhadap layanan tersebut. Meskipun terdapat sebagian kecil pengguna yang menyatakan pandangan positif terhadap Dana, sentimen positif hanya mencapai 8,8%. Dilihat dari *word cloud* yang didominasi oleh kata aplikasi, transaksi, dana, saldo, dan masuk, maka dapat diduga bahwa mayoritas ketidakpuasan pengguna disebabkan oleh masalah teknis aplikasi, kegagalan transaksi, saldo yang tidak akurat, transfer dana yang tidak masuk dan sebagainya. Ini menunjukkan adanya kebutuhan peningkatan layanan pelanggan atau pengelolaan masalah yang efektif untuk menangani keluhan pengguna.

Berdasarkan evaluasi yang dilakukan terhadap algoritma NBC dan SVM untuk melakukan klasifikasi pada dataset yang sama, disimpulkan bahwa keduanya menunjukkan kinerja yang sangat baik. Meskipun SVM memiliki akurasi sempurna dan sedikit lebih tinggi dalam presisi, NBC memiliki sedikit keunggulan dalam recall. Kedua algoritme memiliki nilai *F1-score* yang hampir sama, menunjukkan keseimbangan yang baik antara presisi dan recall. Meskipun terdapat perbedaan kecil dalam kinerja mereka, pemilihan antara SVM dan NBC dapat bergantung pada kebutuhan spesifik dari kasus penggunaan, karakteristik data, dan prioritas yang diberikan pada presisi atau *recall*.

Dana perlu melakukan evaluasi mendalam atas masalah yang dihadapi pengguna dan mengambil langkah-langkah untuk meningkatkan layanan mereka. Ini termasuk perbaikan teknis, peningkatan kualitas layanan pelanggan, dan komunikasi yang lebih efektif dengan pengguna. Dengan melakukan perbaikan yang tepat, Dana dapat meningkatkan kepuasan pengguna dan membangun reputasi yang lebih baik di pasar.

**Daftar Pustaka**

- [1] S. Adiningsih, *Transformasi Ekonomi Berbasis Digital di Indonesia*, PT Gramedia Pustaka Utama, 2019.
- [2] R. Anggraini, "Total Transaksi Dana di Xendit Tembus Rp6,3 Triliun Sepanjang 2022," *Bisnis Indonesia*, 25 Januari 2023. [Online]. Available: <https://finansial.bisnis.com/read/20230125/563/1621513/total-transaksi-dana-di-xendit-tembus-rp63-triliun-sepanjang-2022>. [Accessed 6 Februari 2023].
- [3] D. Mardiansyah, "Dana Catatkan Jumlah Transaksi Harian Mencapai 13,5 Juta Per September 2022," *Kompas Gramedia*, 8 November 2022. [Online]. Available: <https://keuangan.kontan.co.id/news/dana-catatkan-jumlah-transaksi-harian-mencapai-135-juta-per-september-2022>. [Accessed 6 Februari 2023].
- [4] F. A. Larasati, D. E. Ratnawati, dan B. T. Hanggara, "Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Vol. 6 No. 9, pp. 4305-4313, 2022.
- [5] Surohman, F. F. Wati, S. Aji, dan Rousyati, "Analisa Sentimen Terhadap Review Fintech Dengan Metode Naive Bayes Classifier Dan K- Nearest Neighbor," *Jurnal Sains dan Manajemen*, vol. 8 no. 1, pp. 93-105, 2020.
- [6] Y. Harfian, "Klasifikasi Sentimen Aplikasi Dompot Digital Dana Pada Komentar Di Instagram Menggunakan Naive Bayes Classifier," *Tugas Akhir, Universitas Islam Negeri Sultan Syarif Kasim Riau*, 2021.
- [7] A. A. Muhammad, "Analisis Sentimen Pengguna Aplikasi Dana Berdasarkan Ulasan Pada Google Play Store Menggunakan Metode Support Vector Machine," *Skripsi Thesis. Universitas Pembangunan Nasional Veteran Jakarta.*, 2022.
- [8] E. S. Basryah, A. Erfina, dan C. Warman, "Analisis Sentimen Aplikasi Dompot Digital Di Era 4.0 Pada Masa Pandemi Covid-19 Di Play Store Menggunakan Algoritma Naive Bayes Classifier," *SISMATIK (Seminar Nasional Sistem Informasi dan Manajemen Informatika)*, pp. 189-196, 2021.
- [9] B. Gunawan, H. S. Pratiwi, dan E. E. Pratama, "Sistem Analisis Sentimen Pada Ulasan Produk Menggunakan Metode Naive Bayes," *Jurnal Edukasi dan Penelitian Informatika*, vol. 4 no. 2, pp. 113-118, 2018.
- [10] A. Erfina, A. Saepulrohman, D. Lestari, dan E. S. Basryah, "Analisis Sentimen Aplikasi Pembelajaran Online Di Play Store Pada Masa Pandemi Covid-19 Menggunakan Algoritma Support Vector Machine (SVM)," *Seminar Nasional Informatika (SEMASIF)*, pp. 145-152, 2020.
- [11] S. A. F. Ginting, "Analisis Sentimen Dalam Evaluasi Kepuasan Pelanggan Terhadap Pelayanan Jasa Transportasi Online Gojek dan Grab", *Tugas AKhir*, Universitas Sumatra Utara, 2021.
- [12] H. P. Doloksaribu dan Y. T. Samuel, "Komparasi Algoritma Data Mining Untuk Analisis Sentimen Aplikasi Pedulilindungi," *Jurnal Teknologi Informasi*, vol. 16, no. 1, 2022.
- [13] A. Esuli dan F. Sebastiani, *Sentiment Analysis, A Handbook of Theory and Practice*, Springer, 2015.
- [14] M. Ichwan, I. A. Dewi, dan Z. Muharom, "Klasifikasi Support Vector Machine (SVM) Untuk Menentukan Tingkat Kemanisan Mangga Berdasarkan Fitur Warna," *MIND Journal*, vol. 3, no. 2, pp. 16-24, 2018.
- [15] M. A. Ramadhan dan R. Andarsyah, *Klasifikasi Text Spam Menggunakan Metode Support Vector Machine Dan Naive Bayes*, Parongpong, Bandung Barat: Buku Pedia, 2022.
- [16] A. Nurian dan B. N. Sari, "Analisis Sentimen Ulasan Pengguna Aplikasi Google Play Menggunakan Naive Bayes," *Jurnal Informatika Dan Teknik Elektro Terapan (JITET)*, vol. 11, no. 3s1, pp. 829-835, 2023.