

# Sentiment Analysis of DANA Application Reviews on Google Play Store Using Naïve Bayes Classifier Algorithm Based on Information Gain

Cindy Caterine Yolanda, Syafriandi\*, Yenni Kurniawati, Dina Fitria

Departemen Statistika, Universitas Negeri Padang, Padang, Indonesia

\*Corresponding author: [syafriandi\\_math@fmipa.unp.ac.id](mailto:syafriandi_math@fmipa.unp.ac.id)

Submitted : 08 Januari 2024

Revised : 26 Januari 2024

Accepted : 01 Februari 2024

## ABSTRACT

*The DANA application is a digital wallet application that is widely used by Indonesians because of its ease of use and the features offered provide convenience in everyday life. However, there are some shortcomings of this application that cause users to reconsider their use. Reviews from other users are one of the guidelines for someone in making a decision to use this application. Sentiment analysis is a method to categorize user reviews into positive and negative reviews. The purpose of this research is to evaluate the results of sentiment classification of DANA application user reviews on the Google Play Store service using the Naïve Bayes Classifier method and Information Gain feature selection. In addition, this study aims to assess the effect of applying IG feature selection on the performance of the resulting model. In this study, reviews are divided into two categories, namely positive and negative based on lexicon-based labeling. Furthermore, data weighting, feature selection, and data division are carried out with a proportion of 80% train data and 20% test data before model building. There are two models, namely a model without feature selection (NBC model) and a model with feature selection (NBC-IG model). The evaluation results showed that the model without feature selection (NBC model) with 1106 features performed well, with 82.91% accuracy, 83.96% precision, and 90.23% recall. Meanwhile, the model with feature selection (NBC+IG model) with 536 features showed higher performance, with 85.09% accuracy, 85.79% precision, and 92.09% recall. The application of IG feature selection with the IG value limit parameter  $> 0.01$  in the NBC model successfully reduced the number of features by 570, and improved model performance with an increase in accuracy by 2.18%, precision by 1.83%, and recall by 1.86%.*

**Keywords:** DANA application, Information Gain, NBC, Sentiment Analysis.



This is an open access article under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in medium, provided the original work is properly cited. ©2022 by author and Universitas Negeri Padang.

## I. PENDAHULUAN

Kemajuan teknologi telah mencakup segala aspek kehidupan, termasuk dalam proses pembayaran atau transaksi, sehingga masyarakat cenderung beralih ke transaksi digital. Pada tahun 2018 hingga tahun 2023, terjadi peningkatan yang sangat signifikan dalam nilai transaksi uang elektronik. Menurut laporan Bank Indonesia (2023), pada bulan April 2023, nilai transaksi mencapai Rp 37,46 triliun. Angka ini menunjukkan kenaikan sebesar 1,4% dibandingkan dengan bulan sebelumnya (Maret 2023), dan lebih tinggi 5,8% dibandingkan dengan tahun sebelumnya (April 2022).

Peningkatan nilai belanja uang elektronik dapat memicu banyaknya perusahaan rintisan yang mengembangkan inovasi produk di bidang layanan keuangan berbasis teknologi, atau yang dikenal sebagai *Financial Technology (fintech)*. Salah satu *fintech* yang banyak digunakan di Indonesia adalah dompet digital seperti DANA, Link Aja, GoPay dan OVO. DANA menjadi aplikasi keuangan yang paling banyak diunduh di Indonesia. DANA telah digunakan oleh lebih dari 50 juta pengguna di Indonesia dengan rating 4,5 di *Google Play Store*.

*Google Play Store* toko aplikasi resmi untuk semua perangkat berbasis sistem operasi Android. Pada *Google Play Store*, terdapat fitur *rating* dan ulasan aplikasi yang digunakan oleh pengguna untuk memberikan *feedback* terkait aplikasi tersebut. Menurut Saputra dkk. (2019), ulasan telah menjadi alat yang efektif dan efisien untuk memperoleh informasi mengenai suatu produk atau jasa. Di era digital saat ini, banyak pengguna mengandalkan ulasan sebelum membuat keputusan untuk menggunakan suatu produk. Namun, jumlah ulasan yang mencapai ratusan hingga jutaan membuat analisis manual menjadi sulit dilakukan (Masturoh dan Pohan, 2021).

Metode yang digunakan untuk menggali informasi dari teks ulasan dikenal dengan analisis sentimen. Nurjanah dkk. (2017), memaparkan bahwa analisis sentimen bertujuan untuk menilai apakah pandangan atau komentar terkait suatu isu bersifat negatif atau positif, serta mampu digunakan sebagai pedoman dalam peningkatan layanan atau kualitas suatu produk. Dalam melakukan analisis sentimen, terdapat beragam metode

yang dapat diterapkan, dan dalam penelitian ini dipilih metode *Naive Bayes Classifier* (NBC). Pemilihan ini didasarkan pada hasil penelitian oleh Falih dkk. (2022) tentang penerapan seleksi fitur *Information Gain* (IG) dalam analisis sentimen pengguna aplikasi Flip di *Google Play Store* menggunakan algoritma SVM. Hasilnya menunjukkan bahwa IG dapat meningkatkan akurasi model SVM dari 91,97% menjadi 96,25%, mencatat peningkatan sebesar 4,29%.

Berangkat dari uraian di atas, dalam penelitian ini dilakukan *sentiment analysis* terhadap *review* aplikasi DANA di *Google Play Store* dengan metode NBC berbasis seleksi fitur IG. Kedua metode ini telah terbukti memberikan akurasi yang baik pada penelitian sebelumnya. Penelitian ini dilakukan guna mengetahui hasil evaluasi yang diperoleh oleh model yang dibangun yaitu model NBC tanpa seleksi fitur dan model NBC dengan seleksi fitur, serta mengetahui pengaruh penerapan seleksi fitur IG terhadap performa model yang dihasilkan.

## II. METODE PENELITIAN

Dalam penelitian ini, pendekatan analisis yang digunakan ialah *Naive Bayes Classifier* (NBC) dengan penerapan seleksi fitur *Information Gain* (IG). NBC merupakan suatu metode klasifikasi yang menggunakan perhitungan probabilitas. NBC banyak digunakan untuk analisis sentimen karena merupakan metode yang sederhana tetapi memiliki akurasi serta performansi yang tinggi dalam pengklasifikasian teks (Routray, 2013). Seleksi fitur IG digunakan untuk mengurangi jumlah fitur agar dimensi data pada saat klasifikasi tidak terlalu tinggi. Metode ini digunakan karena dapat memilih fitur yang memberikan informasi paling relevan terkait kelas atau label dalam dataset. Proses analisis sentimen dilakukan dengan menggunakan bahasa pemrograman *Python* di *platform Google Colaboratory*. Rangkaian langkah-langkah analisis pada penelitian ini melibatkan beberapa tahapan, seperti yang dijelaskan berikut ini.

### 1. Melakukan pengumpulan data

Data dikumpulkan dari layanan *Google Play Store* dengan teknik *web scraping*, menggunakan bahasa pemrograman *python* dan *library google-play-scrapers* yang harus diinstalasi terlebih dahulu ke dalam *python*. Ulasan yang diperoleh dari proses *scraping* disimpan dalam bentuk *dataframe*, kemudian diubah ke dalam format *csv* dan *xlsx* yang berisi fitur-fitur ulasan.

### 2. Melakukan *text preprocessing*

*Text preprocessing* merupakan langkah-langkah untuk membersihkan data dari atribut-atribut yang tidak dibutuhkan sehingga data bisa diproses pada tahap berikutnya. Merujuk dari penelitian yang dilakukan oleh Muhammad dkk. (2022), langkah-langkah *text preprocessing* yang dilakukan dalam penelitian ini mencakup:

#### a) *Cleaning Data*

*Cleaning* melibatkan proses eliminasi elemen-elemen yang tidak memiliki fungsi dan mengganggu dalam ulasan, seperti tanda baca, emoji, tautan dan angka. Gupta dan Sehgal (2021), memaparkan bahwa tujuan proses ini adalah untuk mengurangi *noise* dan mencegah masalah kualitas data, seperti data tidak lengkap, duplikat, *outlier*, dan tidak valid.

#### b) *Case Folding*

*Case folding* dilakukan dengan tujuan meratakan semua huruf dalam data menjadi *lower case*, agar terhindar dari ketidakkonsistenan seperti *case sensitive* yang tidak seragam (Sari dkk, 2019).

#### c) Tokenisasi

Tokenisasi adalah proses pemisahan teks menggunakan spasi atau tanda baca. *Input* dari proses ini adalah data berupa teks, dan *outputnya* adalah token-token kata (Jo, 2019).

#### d) Normalisasi

Tahap *spelling normalization* atau sering disebut normalisasi kata, merupakan bagian dari proses yang bertujuan untuk melakukan perbaikan terhadap berbagai kata dalam teks dokumen yang mungkin salah eja, disingkat, atau tidak sesuai dengan kaidah bahasa Indonesia. Normalisasi ejaan dilakukan dengan tujuan untuk mengurangi variasi dalam ejaan kata-kata yang sebenarnya memiliki makna yang sama, sehingga dapat membantu menyatukan dan merapikan dimensi kata.

#### e) *Stopword removal*

*Stopword removal* merupakan proses eliminasi kata-kata yang tidak relevan atau tidak memiliki pengaruh dalam sebuah teks. Menurut Muhammad dkk. (2019), *stopword* merujuk pada kata-kata yang sering muncul tetapi tidak berkontribusi pada sentimen, sehingga kata-kata tersebut dianggap sebagai *noise* dan dihapus dalam proses analisis. Korpus *stopword* yang digunakan berasal dari kamus yang disediakan oleh *library NLTK*, kemudian ditambah dengan beberapa *stopword* tambahan.

#### f) *Stemming*

Proses *stemming* digunakan untuk mengubah kata-kata dalam data ulasan menjadi bentuk dasarnya dengan bantuan *library* sastrawi. Cara kerja proses ini adalah dengan menghapus imbuhan atau afiks yang melekat pada kata seperti me-, ber-, di-, -kan, dan -an.

3. Melakukan pelabelan data

Pada langkah ini, dataset akan diartikan terlebih dahulu ke dalam bahasa Inggris karena proses pelabelan data akan menggunakan kamus VADER *sentiment* yang berbahasa Inggris. Data akan diberi label menggunakan metode *lexicon based* dengan menggunakan *library VaderSentiment* pada *python*. Penentuan kelas sentimen bergantung pada hasil perhitungan *compound score* atau skor total dari nilai polaritas kata yang tersimpan di dalam kamus *lexicon*. Perhitungan *compound score* menurut Hutto dan Gilbert (2014), dapat dilakukan menggunakan persamaan:

$$\text{compound score (cs)} = \frac{x}{\sqrt{x^2 + \alpha}} \quad (1)$$

Keterangan :

$x$  = total *score polarity* dari setiap *term* yang terdapat dalam dokumen

$\alpha$  = konstanta normalisasi dengan nilai 15

Proses pemberian label pada data dilakukan untuk menandai setiap ulasan sebagai positif atau negatif. Jika nilai *compound score* lebih besar dari 0 maka ulasan tersebut masuk pada kelas sentimen positif. Jika *compound score* kurang dari 0 maka ulasan tersebut masuk kedalam kelas sentimen negatif.

4. Melakukan pembobotan fitur

Pemberian bobot kata memiliki tujuan untuk memberikan nilai tertentu pada setiap kata atau istilah yang digunakan sebagai masukan dalam proses klasifikasi. Hal ini dilakukan karena model klasifikasi hanya mampu menerima kata-kata dalam bentuk numerik. Dalam penelitian ini, perhitungan bobot kata dilakukan dengan menerapkan metode *Term Frequency – Inverse Document Frequency* (TF-IDF). TF-IDF merupakan metode yang memberikan nilai bobot pada kata-kata berdasarkan frekuensi kemunculan kata tersebut dalam suatu dokumen. *Term Frequency* (TF) menunjukkan frekuensi kemunculan suatu *term*, yang dapat berupa kata, frasa, atau elemen dari indeks pada naskah. Semakin sering suatu *term* muncul, semakin tinggi bobotnya. *Inverse Document Frequency* (IDF) digunakan untuk mengurangi bobot *term* di berbagai dokumen yang sering muncul. Menurut Praptiwi (2018), hal ini dilakukan karena *term* tersebut cenderung menjadi umum dan kurang penting dalam mendeskripsikan suatu dokumen, sebab nilai TF-IDF yang tinggi menandakan *term* tersebut memiliki nilai informasi yang tinggi. Perhitungan TF-IDF menurut Jo (2019) dapat dilakukan menggunakan persamaan berikut.

$$IDF_i = \log_{10} \frac{N}{DF_i} \quad (2)$$

$$w_{i,j} = TF_{i,j} * IDF_i \quad (3)$$

Keterangan:

$w_{i,j}$  = bobot TF-IDF kata  $i$  pada dokumen  $j$ .

$N$  = jumlah total dokumen

$DF_i$  = jumlah dokumen dengan kata  $i$  di dalamnya.

$IDF_i$  = inverse frekuensi dokumen tiap kata

$TF_{i,j}$  = jumlah kemunculan kata  $I$  pada dokumen  $j$ .

5. Melakukan seleksi fitur

Pemilihan fitur adalah proses identifikasi fitur-fitur yang memiliki tingkat kepentingan, signifikansi, relevansi, kegunaan, atau kontribusi yang tinggi dari seluruh kumpulan data, dengan tujuan mengurangi dimensi data. Pemanfaatan metode seleksi fitur dengan cermat tidak hanya berkontribusi pada pemahaman atribut yang signifikan untuk kelas tertentu, tetapi juga berpotensi meningkatkan tingkat akurasi klasifikasi sebagaimana disampaikan oleh Koncz dkk. (2011, 358). *Information Gain* adalah suatu metode pengukuran perubahan nilai entropi setelah dataset mengalami pemisahan berdasarkan fitur tertentu. *Information Gain* mengukur sejauh mana suatu fitur menyediakan informasi yang relevan mengenai kelas atau label pada dataset. Gupta dan Sehgal (2021), menyatakan bahwa entropi adalah ukuran yang mengindikasikan tingkat ketidakteraturan atau ketidakmurnian data dengan mengukur ketidakpastian pada suatu fitur tertentu. Menurut Sari dkk. (2019), rumus yang digunakan dalam proses seleksi fitur *Information Gain* adalah sebagai berikut.

a) Rumus entropi *record* dan total

$$\text{Entropy (S)} = - \sum_{i=1}^m (P_i) \log_2(P_i) \quad (4)$$

b) Rumus entropi fitur

$$Entropy_F(S) = \sum_{j=1}^n -P_j * Ent(S_j) \quad (5)$$

c) Rumus *Information Gain*

$$Gain(S, F) = Entropy(S) - |Entropy_F(S)| \quad (6)$$

Keterangan:

$S$  = sampel (*record* atau total)  
 $F$  = fitur  
 $m$  = total maksimum fitur  
 $n$  = total nilai pada kelas  
 $P_i$  = perbandingan total sampel di kelas  $i$  terhadap total sampel  
 $P_j$  = perbandingan total sampel  $j$  terhadap total sampel pada fitur  
 $Ent(S_j)$  = nilai *entropy* untuk sampel  $j$ .

6. Membangun model klasifikasi sentimen

Sebelum melakukan pemodelan klasifikasi, dataset dibagi menjadi data *training* dan data *testing*. Data *training* digunakan sebagai bahan pembelajaran oleh model dalam melakukan klasifikasi. Sedangkan data *testing* digunakan untuk mengukur kinerja dari model yang telah dibentuk. Jumlah data *training* yang besar dapat menghasilkan akurasi model yang lebih baik. Komposisi pembagian data yang digunakan dalam penelitian ini adalah 80% : 20%.

Pemodelan klasifikasi sentimen dilakukan dengan menggunakan metode *Naïve Bayes Classifier* (NBC). NBC ialah suatu metode pengklasifikasian statistik yang berlandaskan pada teorema Bayes, yaitu menghitung probabilitas untuk memprediksi kejadian masa depan berdasarkan pengalaman sebelumnya (Tangraeni & Melkior, 2022). Secara umum, nilai probabilitas NBC disajikan sebagai berikut.

$$P(V_j|a_i) = \frac{P(a_i|v_j)P(v_j)}{P(a_i)}$$

Keterangan:

$P(V_j|a_i)$  = peluang kategori kelas  $j$  ketika kata ke- $i$  muncul.  
 $P(a_i|v_j)$  = peluang kata ke- $i$  terdapat dalam kategori kelas  $j$   
 $P(v_j)$  = peluang jumlah kejadian kategorik kelas  $j$   
 $P(a_i)$  = peluang munculnya kata ke- $i$ .

Klasifikasi bertujuan untuk menentukan kelas terbaik untuk suatu teks dengan cara mencari probabilitas tertinggi dari semua kategori kelas yang diuji, dengan mencari nilai *Maximum A Posteriori* (MAP) untuk kelas  $V_{MAP}$ , seperti yang dijelaskan dalam Persamaan (6) berikut.

$$V_{MAP} = \underset{v_j \in V}{\arg \max} P(v_j) \times \prod_i P(a_i|v_j) \quad (7)$$

Adapun nilai  $P(v_j)$  yang dihitung pada saat *training* dan nilai  $P(a_i|v_j)$  untuk probabilitas kata  $a_i$  untuk setiap kategori dapat diperoleh dari persamaan (7) dan (8).

$$P(v_j) = \frac{docs_j}{training} \quad (8)$$

$$P(a_i|v_j) = \frac{n_i + 1}{n + kosakata} \quad (9)$$

Dimana  $docs_j$  mewakili jumlah dokumen dalam ketagori  $j$  dan *training* adalah total dokumen yang digunakan selama proses *training*. Sedangkan itu,  $n_i$  mengindikasikan jumlah kemunculan kata  $a_i$  dalam kategori  $v_j$ ,  $n$  adalah jumlah kosakata yang muncul pada kategori  $v_j$  dan kosakata mencakup jumlah kata unik dalam semua data latih (Rodiyansyah & Winarko, 2013).

7. Melakukan evaluasi model

Evaluasi model digunakan untuk menilai kinerja dari model yang telah dikembangkan. Hasil klasifikasi dari data uji akan direpresentasikan dalam bentuk *confusion matrix* yang terdiri dari empat bagian, yaitu *True Positive* (TP), *False Positive* (FP), *False Negative* (FN), dan *True Negative* (TN). Pada *confusion matrix* terdapat beberapa nilai yang dapat dihitung, seperti akurasi, presisi, dan *recall*. Akurasi dapat diartikan sebagai tingkat kesesuaian

antara nilai yang diprediksi dengan nilai sebenarnya, diperoleh dengan membandingkan data yang diklasifikasikan secara benar terhadap keseluruhan dataset. Di sisi lain, presisi didefinisikan sebagai jumlah data teks yang diklasifikasikan dengan benar sebagai positif, dibagi oleh keseluruhan data yang diklasifikasikan sebagai positif. *Recall*, atau disebut juga sebagai *sensitivity*, adalah persentase dari data kategori positif yang dapat diklasifikasikan dengan benar oleh sistem, dan berfungsi untuk menilai seberapa baik sistem dalam mengambil kembali informasi yang relevan. Perhitungan serta tabel pengukuran akurasi menggunakan *confusion matrix* dapat dilihat seperti di bawah ini.

**Tabel 1.** *Confusion Matrix*

Actual Class	Predicted Class	
	Yes	No
	Yes	TP (True Positive)
No	FN (False Positive)	TN (True Negative)

(Sumber: Bramer, 2007).

$$Akurasi = \frac{TN + TP}{TN + TP + FN + FP} \times 100\% \tag{10}$$

$$Presisi = \frac{TP}{TP + FP} \times 100\% \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \tag{12}$$

### III. HASIL DAN PEMBAHASAN

#### A. Pengumpulan Data

Dataset yang digunakan berasal dari ulasan aplikasi DANA yang diambil dari layanan *Google Play Store* menggunakan teknik *scraping* dari *library google-play-scraper*. Proses penambangan data dilakukan pada tanggal 8 September 2023 dengan memasukkan *link* aplikasi DANA yang terdapat di laman *Google Play Store* ke dalam perintah yang telah disusun dalam bahasa pemrograman *python*. Jumlah data yang berhasil diambil sebanyak 1500 ulasan terbaru yang ditulis dalam bahasa Indonesia. Setelah data ulasan diperoleh, langkah selanjutnya adalah mengubahnya ke dalam bentuk *dataframe*, di mana hanya fitur-fitur yang relevan seperti isi ulasan dan *rating* yang dipilih untuk dipertahankan. Kemudian, *dataframe* yang telah disusun tersebut disimpan dalam format *csv* dan *xlsx* agar dapat diolah pada tahap analisis selanjutnya.

#### B. Text Preprocessing

Tahapan awal setelah diperoleh dataset adalah *text preprocessing*, tahapan *text preprocessing* yang dilakukan pada penelitian ini meliputi *cleaning*, *case folding*, *tokenizing*, *normalizing*, *stopword removal*, dan *stemming*. Adapun hasil *text preprocessing* yang dilakukan dapat dilihat pada Tabel 2.

**Tabel 2.** Contoh Tahapan *Text Preprocessing*

Nama Proses	Output
Dataset	Buat admin dana yg terhormat. Sudah 1 tahun lebih pengaduan saya berikan. Tetapi hasilnya ttp nihil. Lewat email sama saja. Syarat2 sdh dipenuhi pun ttp akun dana saya tdk bisa kembali. Apalagi saldo saya. Jawabannya selalu menunggu dan menunggu terus. Menandakan pelayanan sangat lambat 😞
<i>Cleaning</i>	Buat admin dana yg terhormat Sudah tahun lebih pengaduan saya berikan Tetapi hasilnya ttp nihil Lewat email sama saja Syarat sdh dipenuhi pun ttp akun dana saya tdk bisa kembali Apalagi saldo saya Jawabannya selalu menunggu dan menunggu terus Menandakan pelayanan sangat lambat
<i>Case Folding</i>	buat admin dana yg terhormat sudah tahun lebih pengaduan saya berikan tetapi hasilnya ttp nihil lewat email sama saja syarat sdh dipenuhi pun ttp akun dana saya tdk bisa kembali apalagi saldo saya jawabannya selalu menunggu dan menunggu terus menandakan pelayanan sangat lambat
<i>Tokenizing</i>	['buat', 'admin', 'dana', 'yg', 'terhormat', 'sudah', 'tahun', 'lebih', 'pengaduan', 'saya', 'berikan', 'tetapi', 'hasilnya', 'ttp', 'nihil', 'lewat', 'email', 'sama', 'saja', 'syarat', 'sdh', 'dipenuhi', 'pun', 'ttp', 'akun', 'dana', 'saya', 'tdk', 'bisa', 'kembali', 'apalagi', 'saldo', 'saya', 'jawabannya', 'selalu', 'menunggu', 'dan', 'menunggu', 'terus', 'menandakan', 'pelayanan', 'sangat', 'lambat']
<i>Normalizing</i>	['buat', 'admin', 'dana', 'yang', 'terhormat', 'sudah', 'tahun', 'lebih', 'pengaduan', 'saya', 'berikan', 'tetapi', 'hasilnya', 'tetap', 'nihil', 'lewat', 'email', 'sama', 'saja', 'syarat', 'sudah', 'dipenuhi', 'pun',

Nama Proses	Output
	'tetap', 'akun', 'dana', 'saya', 'tidak', 'bisa', 'kembali', 'apalagi', 'saldo', 'saya', 'jawabannya', 'selalu', 'menunggu', 'dan', 'menunggu', 'terus', 'menandakan', 'pelayanan', 'sangat', 'lambat']
<i>Stopword removal</i>	['admin', 'dana', 'terhormat', 'pengaduan', 'hasilnya', 'nihil', 'email', 'syarat', 'dipenuhi', 'akun', 'dana', 'saldo', 'jawabannya', 'menunggu', 'menunggu', 'menandakan', 'pelayanan', 'lambat']
<i>Stemming</i>	['admin', 'dana', 'hormat', 'adu', 'hasil', 'nihil', 'email', 'syarat', 'penuh', 'akun', 'dana', 'saldo', 'jawab', 'tunggu', 'tunggu', 'tanda', 'layan', 'lambat']

Pada saat melakukan *text preprocessing* terdapat beberapa data yang dihapuskan karena berupa data duplikat atau ulasannya menjadi kosong setelah melewati proses *stopword removal*. Jumlah dataset yang tersisa setelah melewati proses *text preprocessing* adalah sebanyak 1372 data. Dataset ini kemudian diterjemahkan ke dalam bahasa Inggris agar dapat dilakukan proses pelabelan.

### C. Pelabelan Data

Langkah berikutnya yaitu pemberian label pada data ulasan yang sudah dipersiapkan. Proses penentuan label dilaksanakan secara otomatis dengan memanfaatkan metode berbasis *lexicon*, dimana sentimen terkategori menjadi dua kelas, yakni positif dan negatif. Kamus *lexicon* yang digunakan dalam penelitian ini merupakan kamus *vader sentiment*. Proses pelabelan menggunakan *lexicon* dilaksanakan dengan menginstal *library Python vadersentiment* pada *Google Colaboratory*. Penentuan kelas sentimen bergantung pada hasil perhitungan *compound score* yang dilakukan menggunakan Persamaan (1). Ulasan dengan nilai *compound score* yang lebih besar dari 0 dikategorikan pada kelas sentimen positif dan yang kurang dari 0 dikategorikan ke dalam kelas sentimen negatif. Hasil dari proses pelabelan dapat ditemukan dalam Tabel 3.

**Tabel 3.** Hasil Pelabelan Data

Kelas Sentimen	Jumlah
Positif	788
Negatif	584

### D. Pembobotan Term Frequency – Inverse Document Frequency

Dataset yang telah melalui proses praproses teks dan pelabelan akan mengikuti tahap pembobotan selanjutnya. Proses ini esensial untuk memberikan nilai bobot pada setiap kata, memungkinkan pengolahan data oleh model *machine learning* yang hanya dapat menerima data numerik. Pembobotan dilakukan menggunakan metode *Term Frequency – Inverse Document Frequency* (TF-IDF). Langkah awal pembobotan dengan metode TF-IDF adalah menghitung *Term Frequency* (TF) dalam setiap dokumen, diikuti dengan perhitungan *Inverse Document Frequency* (IDF) pada setiap kata dalam dokumen menggunakan Persamaan (2). Setelah memperoleh nilai TF dan IDF, langkah berikutnya adalah menghitung bobot TF-IDF pada setiap kata menggunakan Persamaan (3). Hasil pembobotan kata dengan metode TF-IDF, disajikan dalam bentuk *sparse* matriks, dapat ditemukan pada Tabel 4.

**Tabel 4.** Matriks Hasil Pembobotan TF-IDF

Term No. Dok	adil	admin	adu	agen	...	warung	wib	wifi	zonk
1.	0	0,297088	0,358122	0	...	0	0	0	0
2.	0	0	0	1,119687	...	0	0	0	0
3.	0	0	0	0	...	0	0	0	0
4.	0	0	0	0	...	0	0	1,343625	0
...	...	...	...	...	...	...	...	...	...

### E. Seleksi Fitur Information Gain

Proses pemodelan klasifikasi NBC yang didasarkan pada *Information Gain* melibatkan langkah seleksi fitur *Information Gain*. Perhitungan nilai *Information Gain* (IG) untuk setiap fitur dilakukan dengan menggunakan Persamaan (4), (5), dan (6). Fitur yang terpilih adalah fitur yang memiliki nilai IG lebih besar dari 0,01. Setelah melalui tahap seleksi fitur, dataset awal yang awalnya memiliki 1106 fitur telah direduksi menjadi 536 fitur, sedangkan 570 fitur lainnya dihapuskan karena memiliki nilai IG di bawah 0,01. Dataset hasil seleksi fitur ini kemudian digunakan untuk pemodelan dengan metode NBC, dan hasilnya dibandingkan dengan pemodelan NBC tanpa seleksi fitur.

### F. Klasifikasi Naïve Bayes Classifiers

Sebelum melakukan proses pemodelan dilakukan *splitting* dataset menjadi 80% *data training* dan 20% *data testing* dari 1372 data. Pengklasifikasian dilakukan dengan metode NBC terhadap 1097 *data training*. Pada

kelompok data ini, dihitung probabilitas dari masing-masing kategori kelas menggunakan persamaan (8) dan probabilitas kata  $a_i$  untuk setiap kategori sentimen dengan persamaan (9). Model yang telah dibangun kemudian akan diaplikasikan pada 275 *data testing*. Untuk memprediksi kelas dari data, dihitung nilai probabilitas tertinggi ( $V_{MAP}$ ) dari kategori sentimen yang diujikan dengan Persamaan (7). Tabel 5 berikut menyajikan contoh hasil pengujian menggunakan metode NBC.

**Tabel 5.** Probabilitas Klasifikasi pada Ulasan Aplikasi DANA

No. Dok	Ulasan	Positif	Negatif	Label
1.	Aplikasi dana bagus semua fiturnya dan mudah di mengerti, tapi satu kendala nya tidak bisa mengganti nama asli.	0,91233	0,08767	Positif
2.	makin update bukannya makin bagus malah makin jelek, tiap buka pasti kalo gak force close harus dibuka ulang lagi apk dana nya	0,472203	0,527797	Negatif

Berdasarkan Tabel 5 terlihat bahwa nilai probabilitas kelas positif pada dokumen 1 sebesar 0,91233 lebih besar dari probabilitas kelas negatif dengan nilai 0,08767 sehingga dokumen 1 diklasifikasikan ke dalam kelas positif. Sementara itu, pada dokumen 2 nilai probabilitas kelas negatif lebih besar daripada kelas positif sehingga dokumen 2 diklasifikasikan berlabel negatif.

### G. Evaluasi Model

Evaluasi telah dilaksanakan pada 275 data uji dengan memanfaatkan hasil dari pemodelan dalam proses klasifikasi, yang diwakili dalam bentuk *confusion matrix*. Tabel 6 berikut menyajikan *confusion matrix* dari model NBC tanpa seleksi fitur dan model NBC dengan seleksi fitur IG.

**Tabel 6.** *Confusion Matrix*

		Model NBC			Model NBC + IG		
		<i>Predicted Class</i>			<i>Predicted Class</i>		
<i>Actual Class</i>		<i>Yes</i>	<i>No</i>		<i>Yes</i>	<i>No</i>	
	<i>Yes</i>	157	17	<i>Yes</i>	163	14	
	<i>No</i>	30	71	<i>No</i>	27	71	

Kinerja model klasifikasi yang telah disusun dapat dievaluasi melalui perhitungan nilai akurasi, presisi, dan *recall* dengan menggunakan Persamaan (10), (11), dan (12). Hasil evaluasi terhadap model yang telah dibentuk tertera dalam Tabel 7 di bawah ini.

**Tabel 7.** Hasil Evaluasi Model

Model	Akurasi	Presisi	Recall
NBC	82,91%	83,96%	90,23%
NBC - IG	85,09%	85,79%	92,09%

Hasil performa model NBC dengan seleksi fitur IG mengalami peningkatan daripada model NBC tanpa IG dilihat dari kenaikan nilai akurasi dari 82,91% menjadi 85,09%, nilai presisi dari 83,96% menjadi 85,79%, dan nilai *recall* dari 90,23% menjadi 92,09%. Hal ini menunjukkan bahwa penerapan seleksi fitur IG dapat meningkatkan hasil evaluasi model NBC dalam mengklasifikasikan sentimen pengguna aplikasi DANA di *Google Play Store*.

## IV. KESIMPULAN

Berdasarkan analisis yang telah dilakukan, hasil evaluasi menunjukkan bahwa model tanpa seleksi fitur (model NBC) dengan jumlah fitur sebanyak 1106 memperoleh performa yang baik, dengan akurasi sebesar 82,91%, presisi sebesar 83,96%, dan *recall* sebesar 90,23%. Sementara model dengan seleksi fitur (model NBC+IG) dengan jumlah fitur 536 menunjukkan performa yang lebih tinggi, dengan akurasi sebesar 85,09%, presisi sebesar 85,79%, dan *recall* sebesar 92,09%. Penerapan seleksi fitur IG dengan parameter batas nilai  $IG > 0,01$  pada model NBC berhasil mengurangi jumlah fitur sebanyak 570, serta meningkatkan performa model dengan peningkatan akurasi sebesar 2,18%, presisi sebesar 1,83%, dan *recall* sebesar 1,86%. Sehingga dapat dikatakan bahwa penerapan seleksi fitur terbukti mampu meningkatkan performa dari model yang dibangun. Untuk penelitian selanjutnya disarankan melakukan kombinasi seleksi fitur Information Gain dengan metode klasifikasi lainnya serta menggunakan dataset ulasan aplikasi DANA terbaru atau aplikasi serupa.

## DAFTAR PUSTAKA

- Bank Indonesia. (2023). *Statistik Sistem Pembayaran dan Infrastruktur Pasar Keuangan (SPIP) Mei 2023*. Jakarta.
- Bramer, M. (2007). *Principle of Data Mining*. London: Springer.
- Falih, I. M., Matondang, N. H., & Chamidah, N. (2022). "Seleksi Fitur Information Gain pada Analisis Sentimen Terhadap Ulasan Aplikasi Flip dengan Algoritma Support Vector Machine", *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, hal. 317-326.
- Herlinawati, N., Yuri, Y., Siti, F., & Windu, G. S. (2020). Analisis Sentimen Zoom Cloud di Play Store Menggunakan Naïve Bayes dan Support Vector Machine. *Journal of Computer Engineering System and Science*. Vol. 5, No. 2.
- Hutto, C. J., & Gilbert, E. (2021). "VADER: A parsimonious rule-based model for sentiment analysis of social media text". *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM*, 216–225.
- Jo, Taeho. (2019). *Text Mining: Concepts, Implementation, and Big Data Challenge*. Berlin: Springer International Publishing.
- Kamalia, A. Z., Zaroni, A. A., & Wangsadanureja, M. (2022). "Analisis Sentimen Pada Ulasan Aplikasi Bibit di Play Store dengan Metode Naïve Bayes, Support Vector Machine, C4.5 Dan K-Nearest Neighbor", *SIGMA: Jurnal Teknologi Pelita Bangsa*, Vol. 13, No. 1, hal. 12-20.
- Masturoh, S., & Pohan, A. B. (2021). Sentiment Analysis Againts the DANA E-Wallet on Google Play Reviews Using the K-Nearest Neighbor Algorithm. *Jurnal PILARI Nusa Mandiri*, Vol. 17, No. 1, hal. 53 – 58.
- Muhammad, A. A., Ermatita, & Prasvita, D. S. (2022). Analisis Sentimen Pengguna Aplikasi Dana Berdasarkan Ulasan Pada Google Play Menggunakan Metode Support Vector Machine. *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, hal. 194 – 204.
- Nurjanah, W. E., Perdana, R. S., & Fauzi, M. A. (2017). "Analisis Sentimen Terhadap Tayangan Televisi Berdasarkan Opini Masyarakat pada Media Sosial Twitter menggunakan Metode K-Nearest Neighbor dan Pembobotan Jumlah Retweet", *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Vol. 1, No. 12, hal. 1750 – 1757.
- P. Gupta & N. K. Sehgal. (2021). *Introduction to Machine Learning in the Cloud with Python: Concepts and Practices*. Springer International Publishing.
- P. Koncz and J. Paralic. (2011). An approach to feature selection for sentiment analysis. *2011 15th IEEE Int. Conf. Intell. Eng. Syst.*, hal. 357–362.
- Praptiwi, D. Y. (2018). *Analisis Sentimen Online Review Pengguna E-Commerce Menggunakan Metode Support Vector Machine dan Maximum Entropy*. Skripsi FMIPA, Universitas Islam Indonesia.
- Rodiyansyah, S. F., & Winarko, E. (2012). "Klasifikasi Posting Twitter Kemacetan Lalu Lintas Kota Bandung Menggunakan Naïve Bayesian Classification", *Indonesia Journal of Computing and Cybernetics Systems (IJCCS)*. Vol. 6, No. 1, Hal. 91–100.
- Routray, P., Swain, C. K. & Mishra, S. P. (2013). "A Survey on Sentiment Analysis". *International Journal of Computer Applications*, Vol. 70, No. 10, Hal. 1-8.
- Saputra, S. A., & Rosiyadi, D. (2019). Analisis Sentimen E-Wallet Pada Google Play Menggunakan Algoritma Naive Bayes Berbasis Particle Swarm Optimization, *Jurnal RESTI: Rekayasa Sistem dan Teknologi Informasi*, Vol. 3, No. 3, hal. 377 – 382.
- Sari, A. E., Widowati, S., & Lhaksana, K. M. (2019). Klasifikasi Ulasan Pengguna Aplikasi Mandiri Online di Google Play Store dengan Menggunakan Metode Information Gain dan Naive Bayes Classifier, *E-Proceeding of Engineering*. Vol. 6, No. 2, hal. 9143 – 9157.
- Tanggaeni, A. I., & Melkior, N. N. (2022). Analisis Sentimen Pengguna E-Government Pada Play Store Menggunakan NBC. *Jurnal Teknik Informatika dan Sistem Informasi*. Vol. 9, No. 2, hal. 785-795.