
**Analisis Sentimen dengan Metode Klasifikasi Naïve Bayes
Dan Seleksi Fitur Information Gain
(Studi Kasus: Ulasan Aplikasi PeduliLindungi)**

Natalia Syafitri Kustanto¹, Nurul Gusriani², Firdaniza³

^{1,2,3} Program Studi S-1 Matematika, Fakultas MIPA, Universitas Padjadjaran

Jl. Raya Bandung Sumedang KM 21 Jatinangor Sumedang 45363

Email : natalia18001@mail.unpad.ac.id, nurul.gusriani@unpad.ac.id, firdaniza@unpad.ac.id

Abstrak

Analisis sentimen merupakan pemrosesan bahasa alami yang bertujuan untuk menemukan opini, mengidentifikasi sentimen apa yang diungkapkan, dan mengklasifikasikannya berdasarkan nilai yang terkandung. Teknologi pembelajaran mesin, seperti Klasifikasi Naïve Bayes banyak digunakan dalam klasifikasi sentimen karena memiliki tingkat akurasi dan kecepatan yang tinggi. Pembelajaran mesin dapat dipadukan dengan metode seleksi fitur untuk meningkatkan efisiensi model. Salah satu metode seleksi fitur adalah metode *Information Gain* yang dapat digunakan untuk menentukan atribut paling efektif. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap aplikasi PeduliLindungi dan menghitung evaluasi performa model berdasarkan nilai metrik *F1-Score*. Pada penelitian ini analisis sentimen dilakukan menggunakan metode Klasifikasi Naïve Bayes dengan penggabungan metode seleksi fitur *Information Gain* untuk meningkatkan performa model dan mempercepat komputasi. Hasil analisis sentimen menggunakan metode klasifikasi Naïve Bayes dan seleksi fitur *Information Gain* menunjukkan bahwa pengguna cenderung memberikan ulasan positif terhadap aplikasi PeduliLindungi. Performa model untuk analisis sentimen pengguna terhadap aplikasi PeduliLindungi ditentukan berdasarkan rata-rata *F1-Score* dengan *10-fold cross validation* diperoleh sebesar 95.1%; artinya model yang diperoleh mempunyai presisi dan *recall* yang baik.

Kata Kunci : analisis sentimen; Naïve Bayes; *Information Gain*; aplikasi PeduliLindungi.

Abstract

Sentiment analysis is a natural language processing that aims to find opinions, identify sentiments expressed, and classify them based on the values they contain. Machine learning technology, such as Naïve Bayes Classification is widely used in sentiment analysis because of its high accuracy and speed. A machine learning method can be combined with feature selection methods to improve model efficiency. One of the feature selection methods is the Information Gain method which can be used to determine the most effective attributes. This research aims to analyze public sentiment towards the PeduliLindungi application and calculate the model's performance based on the F1-score metric value. In this research, sentiment analysis was carried out using the Naïve Bayes Classification method combined with Information Gain Feature Selection Method to improve model performance and speed up the computing process. The results of sentiment analysis using the Naïve Bayes classification method and the Information Gain Selection Feature indicate that users tend to give positive reviews to the PeduliLindungi application. The performance of the model for analyzing user sentiment towards the PeduliLindungi application determined based on an average F1-Score with 10-fold cross-validation was obtained at 95.1%; this means that the model obtained has good precision and recall.

Keywords: *sentiment analysis; Naïve Bayes; Information Gain; PeduliLindungi application.*

1. PENDAHULUAN

Sudah lebih dari kurun waktu dua tahun wabah virus COVID-19 telah menjadi momok permasalahan di seluruh dunia dan menjadi bahan fokus utama bagi para pemimpin dunia dalam mencari cara untuk menghadapi permasalahan ini. Pandemi global yang disebabkan oleh virus yang muncul pertama kali di Kota Wuhan, Provinsi Hubei, China memberikan dampak dan pengaruh yang besar bagi berbagai sektor kehidupan di Indonesia, termasuk di bidang ekonomi, pendidikan, kesehatan, bahkan sosial (Fauziyyah dan Ersyafdi, 2021).

Menurut (KOMINFO, 2021), pemerintah menginisiasi pengembangan aplikasi PeduliLindungi dalam rangka mengoptimalkan perlindungan kesehatan masyarakat dan meminta partisipasi warga untuk mengunduh dan memanfaatkan teknologi tersebut. Pemerintah melalui Kemenkominfo akan terus meningkatkan performa dari PeduliLindungi agar masyarakat tidak menemukan kendala penggunaan, dapat mempermudah serta memberikan rasa aman bagi masyarakat ketika beraktivitas (KOMINFO, 2021). Seiring dengan anjuran pemerintah untuk menggunakan aplikasi ini, beragam tanggapan dari masyarakat terhadap aplikasi PeduliLindungi muncul. Beberapa penelitian telah digunakan untuk melihat persepsi pengguna terhadap aplikasi PeduliLindungi. Kurniawati *et al.*, (2020) melakukan survei kepada 115 pengguna aplikasi PeduliLindungi dan mengukur menggunakan *Technology Acceptance Model* (TAM), hasilnya pengguna dari aplikasi PeduliLindungi akan meningkat karena kemudahan penggunaan dari aplikasi tersebut. Haerani dan Rahmatulloh (2021) menganalisis *User Experience* dengan melakukan survei kepada 420 pengguna, hasil penelitian dari enam skala: daya tarik, kejelasan, efisiensi, ketepatan, stimulasi dan kebaruan masih menghasilkan nilai yang sangat rendah dengan kesimpulan bahwa aplikasi PeduliLindungi belum mendukung program pemerintah secara optimal. Kedua hasil penelitian tersebut memberikan masukan yang baik untuk pengembangan aplikasi PeduliLindungi, namun penggunaan survei

membutuhkan waktu yang lama dan sumber daya untuk melakukannya. Respons dari pengguna aplikasi terbatas sesuai dengan pertanyaan yang ditanyakan oleh peneliti. Perlu dilakukan analisis kembali untuk mendapatkan ulasan dari lebih banyak pengguna agar mendapatkan hasil informasi yang lebih banyak, namun dengan meminimalisir penggunaan waktu dan sumber daya.

Analisis sentimen merupakan kombinasi *text mining* dengan *natural language processing* yang bertujuan untuk menemukan opini, mengidentifikasi sentimen apa yang diungkapkan, dan kemudian mengklasifikasikannya berdasarkan nilai yang terkandung di dalamnya (Liu, 2012). Teknologi pembelajaran mesin (*machine learning*), seperti klasifikasi Naïve Bayes banyak digunakan dalam klasifikasi sentimen karena metode Klasifikasi Naïve Bayes memiliki tingkat akurasi dan kecepatan yang tinggi (Larose, 2006). Penelitian sebelumnya yang menggunakan metode klasifikasi Naïve Bayes adalah untuk menganalisis ulasan dari cuplikan film (Novendri *et al.*, 2020) dengan hasil presisi dan *recall* model secara berurutan sebesar 74.83% dan 75.22%. Penggunaan model tersebut dapat menganalisis data sebanyak 998 ulasan dan memberikan kesimpulan bahwa ulasan mengenai film tersebut cenderung positif sehingga dapat disimpulkan bahwa film tersebut diminati.

Penggunaan metode pembelajaran mesin dapat dipadukan dengan metode seleksi fitur untuk mereduksi data dan meningkatkan efisiensi model. Kohavi & John dalam (Iqbal *et al.*, 2020) membandingkan metode seleksi fitur, hasilnya metode statistik *Chi Square* dan *Information Gain* adalah metode yang paling efektif untuk mengoptimalkan hasil klasifikasi. Forman dalam (Jović, Brkić dan Bogunović, 2015) juga melakukan perbandingan terhadap dua belas metode seleksi fitur untuk klasifikasi teks. Hasilnya metode dengan performa terbaik dari segi presisi untuk klasifikasi teks adalah metode *Information Gain*.

Berdasarkan hal tersebut, penelitian ini menggunakan Klasifikasi Naïve Bayes dengan tujuan untuk mengetahui persepsi pengguna terhadap kualitas layanan Aplikasi

PeduliLindungi. Pada penelitian ini dilakukan penggabungan metode Klasifikasi Naïve Bayes dengan seleksi fitur *Information Gain* untuk meningkatkan performa model dan mempercepat komputasi.

2. KAJIAN PUSTAKA

2.1 Seleksi fitur *Information Gain*.

Seleksi fitur *Information Gain* dilakukan dengan mengukur seberapa efektif suatu atribut tersebut dalam mengklasifikasikan sebuah kelas (Firmahsyah dan Gantini, 2016). Penentuan atribut atau kata terbaik dilakukan dengan menghitung nilai entropi. Nilai entropi dihitung dengan persamaan (1),

$$Entropy(C) = -\sum_{i=1}^m P(C_i) \log P(C_i), \quad (1)$$

dengan C adalah himpunan seluruh ulasan yang dipartisi berdasarkan kelas klasifikasi, dinotasikan $C = \{C_1, C_2, \dots, C_m\}$, m adalah banyaknya kelas klasifikasi, dan $P(C_i)$ adalah peluang dari kelas klasifikasi. Setelah mendapatkan nilai entropi, perhitungan *Information Gain* dapat dilakukan dengan menggunakan persamaan (2) (Firmahsyah dan Gantini, 2016).

$$\begin{aligned} Gain(C, w_k) &= Entropy(C) \\ &- \sum_{f \in F} \frac{n(C_{w_k, f})}{n(C)} Entropy(C_{w_k, f}), \end{aligned} \quad (2)$$

dengan

- $n(C)$: banyaknya seluruh ulasan
- W : himpunan kata yang muncul setidaknya sekali pada ulasan C , $W = \{w_1, w_2, w_3, \dots, w_q\}$
- q : banyaknya kata yang muncul setidaknya sekali pada seluruh ulasan
- f : banyaknya kemunculan kata w_k pada ulasan
- F : himpunan nilai f
- $C_{w_k, f}$: himpunan ulasan dengan kata w_k muncul sebanyak f
- $n(C_{w_k, f})$: banyaknya ulasan dengan kata w_k muncul sebanyak f

2.2 Metode Klasifikasi Naïve Bayes.

Klasifikasi Naïve Bayes adalah metode klasifikasi yang berdasar pada teorema Bayes. Klasifikasi Naïve Bayes bersifat sederhana, walaupun demikian klasifikasi ini memberikan hasil yang baik dan juga memberikan kecepatan dalam memproses data dalam kuantitas yang besar (Aggarwal, 2014). Klasifikasi Naïve Bayes mengasumsikan bahwa ada atau tidaknya sebuah fitur dalam sebuah kelas adalah independen (tidak bergantung), artinya sebuah fitur dalam sebuah kelas tidak punya keterkaitan dengan keberadaan fitur lain dari data yang sama (Jurafsky dan Martin, 2019).

Menurut (Liu *et al.*, 2013), proses Klasifikasi Naïve Bayes adalah sebagai berikut.

- a. Berdasarkan kelas klasifikasinya, misalkan C adalah himpunan seluruh ulasan yang dipartisi menjadi dua kelas klasifikasi yaitu positif dan negatif, dinotasikan $C = \{C_1, C_2\}$. Probabilitas prior masing-masing kelas C dapat dihitung menggunakan persamaan (3).

$$P(C_i) = \frac{n(C_i)}{n(C)} \quad (3)$$

- b. Peluang munculnya setiap kata pada masing-masing kelas klasifikasi dapat dihitung menggunakan persamaan (4).

$$P(W|C_i) = \frac{P(W \cap C_i)}{P(C_i)}. \quad (4)$$

Untuk menghindari nilai $P(W|C_i) = 0$ dilakukan *laplace smoothing* yaitu penambahan suatu nilai, misalkan parameter α , pada setiap perhitungan banyaknya kemunculan kata. Jadi, peluang munculnya setiap kata pada masing-masing kelas klasifikasi adalah

$$P_\alpha(W|C_i) = \frac{P(W \cap C_i) + \alpha}{P(C_i) + q\alpha}. \quad (5)$$

Pendekatan paling sederhana adalah menambahkan satu nilai ($\alpha = 1$) ke setiap jumlah peristiwa yang diamati termasuk kemungkinan hitungan nol.

- c. Misalkan D adalah himpunan ulasan dengan $D = \{d_1, d_2, \dots, d_{n(C)}\}$. Berdasarkan Walpole *et al.*, (2016), untuk menentukan peluang ulasan D termasuk kelas klasifikasi C dapat dihitung dengan teorema Bayes pada persamaan (6)

$$P(C_i|D) = \frac{P(C_i \cap D)}{\sum_{i=1}^m P(C_i \cap D)}$$

$$= \frac{P(C_i)P(D|C_i)}{\sum_{i=1}^m P(C_i)P(D|C_i)} \quad (6)$$

Karena

$$P(D) = P[(C_1 \cap D) \cup (C_2 \cap D) \cup \dots \cup (C_k \cap D)]$$

$$= P(C_1 \cap D) + P(C_2 \cap D) + \dots + P(C_k \cap D)$$

$$= \sum_{i=1}^k P(C_i \cap D)$$

$$= \sum_{i=1}^k P(C_i)P(D|C_i),$$

sehingga persamaan (6) ekuivalen dengan persamaan (7)

$$P(C_i|D) = \frac{P(C_i)P(D|C_i)}{P(D)} \quad (7)$$

Berdasarkan Liu *et al.*, (2013), karena $P(D)$ tidak bergantung pada C dan nilai dari d_i tetap untuk ukuran kumpulan data yang diketahui, $P(D)$ adalah konstan, sehingga penyebut dari persamaan tersebut dapat tidak dihitung. Naïve Bayes mengasumsikan bahwa setiap kata dalam ulasan adalah saling bebas. Oleh karena itu, persamaan tersebut dapat ditulis menjadi

$$P(C_i|D) \propto P(C_i) \prod_{k=1}^{u_{d_i}} [P_{\alpha}(W_k|C_i)]^{f_k}, \quad (8)$$

dengan u_{d_i} adalah banyaknya kata yang muncul setidaknya sekali pada ulasan d_i , dan f_k adalah banyaknya kemunculan kata w_k pada ulasan d_i . Untuk menghindari hasil dengan nilai yang sangat kecil, digunakan transformasi logaritma sehingga dapat ditulis sebagai persamaan (9)

$$\log P(C_i|D) \propto \log P(C_i) \sum_{k=1}^{u_{d_i}} [f_k \log P_{\alpha}(W_k|C_i)]. \quad (9)$$

- d. Kelas dari ulasan d_i ditentukan sebagai kelas C^* , diperoleh dengan menggunakan persamaan (10), yaitu

$$C^* = \max_{i=\{1,2\}} \left\{ \log P(C_i) \sum_{k=1}^{n_{d_i}} [f_k \log P_{\alpha}(W_k|C_i)] \right\}. \quad (10)$$

2.3 Confusion Matrix

Confusion matrix atau *error matrix* adalah tabel matriks yang menampilkan deskripsi kinerja model klasifikasi pada rangkaian data uji (*testing*) yang nilai sebenarnya telah diketahui (Bird, Klein dan Loper, 2009). Elemen dari *confusion matrix* dapat dilihat pada Tabel 1.

Tabel 1. *Confusion matrix*

	Aktual positif	Aktual negatif
Prediksi positif	<i>True Positive</i> (TP)	<i>False Positive</i> (FP)
Prediksi negatif	<i>False Negative</i> (FN)	<i>True Negative</i> (TN)

Confusion matrix digunakan untuk menghitung beberapa metrik dalam klasifikasi, diantaranya adalah akurasi, presisi, *recall*, dan *F1-Score*. Menurut Sokolova *et al.*, dalam (Tharwat, 2020), perhitungan akurasi, presisi, *recall*, dan *F1-Score* adalah sebagai berikut:

- (i) Akurasi

Akurasi merupakan rasio prediksi benar dari jumlah elemen diagonal terhadap jumlah keseluruhan elemen matriks. Akurasi hanya cocok jika digunakan pada saat perbandingan jumlah label data relatif sama (*balanced data*). Secara matematis akurasi dapat dihitung menggunakan persamaan (11),

$$Akurasi = \frac{TP + TN}{TP + FP + FN + TN} \quad (11)$$

- (ii) Presisi

Presisi dapat didefinisikan sebagai derajat reliabilitas model ketika model menghasilkan prediksi positif. Secara matematis presisi dihitung menggunakan persamaan (12),

$$\text{Presisi} = \frac{TP}{TP + FP} \quad (12)$$

(iii) *Recall*

Recall merupakan perbandingan prediksi benar positif dengan keseluruhan data yang benar positif ($TP + FN$). Secara matematis nilai *recall* dihitung menggunakan persamaan (13),

$$\text{Recall} = \frac{TP}{TP + FN} \quad (13)$$

(iv) *F1-Score*

F1-Score menggabungkan presisi dan *recall* dengan mengambil rata-rata harmonik dari keduanya. Nilai *F-Score* berkisar dari nol hingga satu dengan nilai yang tinggi menunjukkan kinerja klasifikasi yang tinggi. Secara matematis nilai *F1-Score* dihitung dengan persamaan (14),

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (14)$$

Metrik *F1-Score* lebih cocok digunakan ketika memiliki data yang tidak seimbang (*imbalanced data*) karena dalam perhitungannya mempertimbangkan jenis kesalahan (*False Positive* atau *False Negative*) dari hasil klasifikasi model.

2.4 K-Fold Cross Validation

Berdasarkan Pitria (2014), *K-Fold Cross Validation* merupakan salah satu metode yang digunakan untuk mengetahui rata-rata keberhasilan dari suatu sistem dengan cara melakukan perulangan dengan mengacak atribut masukan sehingga sistem tersebut teruji untuk beberapa atribut input yang acak. *K-Fold Cross Validation* diawali dengan membagi data sejumlah *k-fold* yang diinginkan. Dalam proses *cross validation* data akan dibagi dalam *k* buah partisi dengan ukuran yang sama $D_1, D_2, D_3, \dots, D_k$ selanjutnya proses pengujian dan pelatihan dilakukan sebanyak *k* kali. Dalam iterasi ke-*i* partisi D_i akan menjadi data uji dan sisanya akan menjadi data latih.

Menurut Kohavi dalam Pitria (2014) menyatakan bahwa penggunaan jumlah *fold* terbaik untuk uji validitas yang dianjurkan adalah *10-fold cross validation*. Pengujian menggunakan *k-fold* ini bertujuan untuk mengetahui rata-rata performa metode Klasifikasi Naïve Bayes yang diterapkan pada analisis sentimen jika diuji dengan data latih dan data uji yang berbeda

3. METODE PENELITIAN

Metode penelitian yang digunakan adalah Klasifikasi Naïve Bayes untuk mengklasifikasi sentimen dari ulasan pengguna aplikasi PeduliLindungi dengan seleksi fitur *Information Gain*.

Adapun langkah dalam penelitian ini adalah sebagai berikut:

1) Pengumpulan data

Data ulasan aplikasi PeduliLindungi didapat dari *Google Play Store* menggunakan bantuan bahasa pemrograman Python dengan *library google-play-scraper*.

2) Pelabelan Data

Data ulasan dengan periode Maret hingga Mei 2022 dilabeli berdasarkan sentimennya (positif atau negatif) sebagai data yang digunakan untuk tahap latih dan uji model klasifikasi Naïve Bayes. Proses pelabelan dilakukan secara manual dengan membaca ulasan satu persatu dan dilakukan oleh tiga orang untuk meminimalisir penilaian yang subjektif.

3) Pra-Proses Data

Proses yang dilakukan dalam pra-proses adalah sebagai berikut:

a. Case Folding

Pada tahap *case folding*, semua huruf dalam dokumen diubah menjadi huruf kecil dan karakter selain huruf dihapus.

b. Tokenizing

Pada tahap *tokenizing*, dilakukan pemisahan kata pada ulasan.

c. Normalisasi

Pada tahap normalisasi, dilakukan perubahan kata tidak baku pada ulasan menjadi kata baku. Proses normalisasi dilakukan menggunakan set data *colloquial-indonesian-lexicon* (Aliyah *et al.*, 2018).

d. Filtering

Pada tahap *filtering*, dilakukan penghapusan *stopword* Bahasa Indonesia menggunakan set data *stopwords-id-satya* (Nugraha dan Chandra, 2017).

e. Stemming

Pada tahap *stemming*, dilakukan proses menghilangkan infleksi kata ke bentuk dasarnya.

4) Pemisahan Data

Merujuk pada penelitian Novendri *et al.* (2020), data ulasan pada penelitian ini dibagi menjadi dua, yaitu:

- Data latih, yaitu data yang digunakan untuk proses pembuatan model klasifikasi, sebanyak 80% dari data ulasan.
- Data uji, yaitu data yang digunakan untuk proses klasifikasi dan pengukuran performa model, sebanyak 20% dari data ulasan.

5) Seleksi fitur *Information Gain*.

Seleksi fitur *Information Gain* dilakukan pada data latih dengan mengukur seberapa efektif setiap kata dalam mengklasifikasikan sebuah kelas. Penentuan atribut atau kata terbaik dilakukan dengan menghitung nilai entropi dan *Information Gain* menggunakan persamaan (1) dan (2).

6) Pemodelan dengan metode klasifikasi Naïve Bayes.

Setelah dilakukan seleksi fitur *Information Gain*, dilakukan pemodelan data latih dengan metode klasifikasi Naïve Bayes menggunakan persamaan (3)-(5).

7) Menentukan kelas klasifikasi dari data uji

Langkah selanjutnya adalah menentukan kelas klasifikasi data uji menggunakan model yang diperoleh dari data latih. Penentuan kelas

klasifikasi dilakukan menggunakan persamaan (6)-(10).

8) Evaluasi model klasifikasi

Evaluasi dilakukan pada seluruh data menggunakan *10-fold cross validation* dengan menghitung nilai akurasi, presisi, *recall*, dan *F1-Score* di setiap iterasi berdasarkan persamaan (11)-(14)

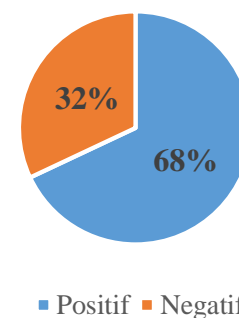
4. HASIL DAN PEMBAHASAN

4.1 Data Penelitian

Data yang digunakan pada penelitian ini adalah data ulasan pengguna aplikasi PeduliLindungi dari bulan Maret hingga Mei 2022 yang didapat dari layanan distribusi aplikasi digital *Google Play Store* (<https://play.google.com/store/apps/details?id=com.telkom.tracencare>). Diperoleh sebanyak 5.162 ulasan aplikasi PeduliLindungi yang diambil dengan menggunakan *library google-play-scraper* pada bahasa pemrograman Python.

4.2 Pelabelan Data

Data ulasan pengguna aplikasi PeduliLindungi dilabeli berdasarkan sentimen (positif atau negatif). Proses pelabelan dilakukan secara manual dengan membaca ulasan satu persatu dan dilakukan oleh tiga orang untuk meminimalisir penilaian yang subjektif. Dari total data ulasan berjumlah 5162, dihasilkan 3510 ulasan dengan sentimen positif dan 1652 ulasan dengan sentimen negatif; sehingga distribusi data memiliki kelas klasifikasi yang tidak seimbang. Distribusi sentimen dari data penelitian dapat dilihat pada Gambar 1.



Gambar 1. Presentase ulasan pengguna berdasarkan sentimen

Tabel 2. Hasil pra-proses ulasan pengguna

No	Ulasan	Hasil pra-proses ulasan	Sentimen
1	Oke	'oke'	Positif
2	kode verifikasi tidak masuk,lemot,terkadang error.	'kode', 'verifikasi', 'tidak', 'masuk', 'lambat', 'kadang', 'error'	Negatif
3	Tolong lah buat apk nya bikin ribet doang cuma isi tanggal lahir aja cape capein belum lagi kalo salah isi harus geser lagi ke tahun yg lumayan jauh tolong buat opsi yang ga merepotkan bintang 1 cacat apk nya	'tolong', 'aplikasi', 'ribet', 'doang', 'isi', 'tanggal', 'lahir', 'capek', 'salah', 'isi', 'geser', 'tahun', 'lumayan', 'jauh', 'tolong', 'opsi', 'tidak', 'repot', 'bintang', 'cacat', 'aplikasi'	Negatif
⋮	⋮	⋮	⋮
5160	Bermanfaat	'bermanfaat'	Positif
5161	Bermanfaat sekali	'bermanfaat'	Positif
5162	Ngerespon di wa lama banget	'respon', 'whatsapp', 'lama', 'banget'	Negatif

4.3 Pra-Proses Data

Data yang sudah dilabeli kemudian dilakukan pra-proses. Langkah yang dilakukan dalam pra-proses data secara bertahap adalah *case folding*, *tokenizing*, normalisasi, *filtering*, dan *stemming*. Langkah tersebut dilakukan dengan menggunakan bahasa pemrograman Python. Data ulasan yang telah melewati tahap pra-proses dapat dilihat pada Tabel 2.

4.4 Pemisahan Data

Data ulasan selanjutnya dibagi menjadi sekitar 80% data latih (4129 ulasan) dan 20% data uji (1033 kalimat). Pembagian data ini dilakukan dengan bantuan bahasa pemrograman Python. Data latih kemudian dikelompokkan berdasarkan label sentimennya, sehingga diperoleh sebanyak 2843 ulasan bersentimen positif dan 1286 ulasan bersentimen negatif.

4.5 Seleksi Fitur Information Gain

Nilai entropi dari C dihitung menggunakan persamaan (1). C_1 adalah ulasan dengan sentimen positif dan C_2 adalah ulasan dengan sentimen negatif, sehingga $C = \{C_1, C_2\}$. Pada penelitian ini terdapat dua kelas klasifikasi positif dan negatif, sehingga $m = 2$.

$$Entropy(C) = -\sum_{i=1}^2 P(C_i) \log P(C_i)$$

$$\begin{aligned} &= -P(C_1) \log P(C_1) - P(C_2) \log P(C_2) \\ &= -\frac{n(C_1)}{n(C)} \log \left(\frac{n(C_1)}{n(C)}\right) - \frac{n(C_2)}{n(C)} \log \left(\frac{n(C_2)}{n(C)}\right) \\ &= -\frac{2843}{4129} \log \left(\frac{2843}{4129}\right) - \frac{1286}{4129} \log \left(\frac{1286}{4129}\right) \\ &= 0,89485. \end{aligned}$$

Selanjutnya dihitung nilai *Information Gain* untuk setiap kata menggunakan persamaan (2). Menggunakan bantuan bahasa pemrograman Python diperoleh nilai *Information Gain* untuk 2892 kata yang muncul pada data latih yang dapat dilihat melalui Tabel 3.

Tabel 3. Nilai *Information Gain* setiap kata pada ulasan

No	Kata	Information Gain
1	tidak	0,12779
2	sertifikat	0,08076
3	vaksin	0,06446
4	bagus	0,03390
5	ok	0,02949
6	update	0,02929
7	tanggal	0,02845
⋮	⋮	⋮
2890	beraktifitas	8,01E-07
2891	nasional	8,01E-07
2892	kuat	8,01E-07

Pemilihan kata dilakukan dengan memilih kata dengan nilai *Information Gain* tertinggi. Pada penelitian ini dipilih 500 kata dengan nilai *Information Gain* tertinggi, sehingga jumlah kata yang dipakai untuk membuat model klasifikasi Naïve Bayes adalah sebanyak 500 kata.

4.6 Pemodelan dengan Metode Klasifikasi Naïve Bayes

Pada tahap ini dilakukan pemodelan pada data latih yang telah dilakukan seleksi fitur *Information Gain*.

- a. Menghitung probabilitas prior masing-masing kelas klasifikasi.

Probabilitas prior ulasan pengguna dengan kelas klasifikasi positif dihitung menggunakan persamaan (3) adalah

$$P(\text{"positif"}) = P(C_1) = \frac{n(C_1)}{n(C)} = \frac{2843}{4129} = 0,68854.$$

Probabilitas prior ulasan pengguna dengan kelas klasifikasi negatif adalah

$$P(\text{"negatif"}) = P(C_2) = \frac{n(C_2)}{n(C)} = \frac{1286}{4129} = 0,31146.$$

- b. Menghitung peluang munculnya setiap kata pada masing-masing kelas

Menggunakan bantuan bahasa pemrograman Python, peluang munculnya setiap kata pada masing-masing kelas dapat dihitung dengan persamaan (5). Hasil perhitungan dapat dilihat pada Tabel 4.

Tabel 4. Peluang munculnya setiap kata pada kelas klasifikasi

No	Kata	$P_{\alpha}(w_k C_1)$	$P_{\alpha}(w_k C_2)$
1	tidak	0,00427	0,05107
2	sertifikat	0,01695	0,03629
3	vaksin	0,00826	0,03948
⋮	⋮	⋮	⋮
498	gratis	0,00007	0,00022
499	seharus	0,00007	0,00022
500	nyiksa	0,00007	0,00022

4.7 Penentuan Kelas Klasifikasi Data Uji

Pada tahap ini dilakukan klasifikasi untuk ulasan pengguna pada data uji. Dari hasil pemisahan data diperoleh data uji sebanyak 1032 kalimat, yaitu dimulai dari ulasan nomor 4130 sampai 5162. Kelas klasifikasi ditentukan berdasarkan nilai C^* pada persamaan (9)-(10). Menggunakan bahasa pemrograman Python, diperoleh hasil klasifikasi data uji yang dapat dilihat pada Tabel 5.

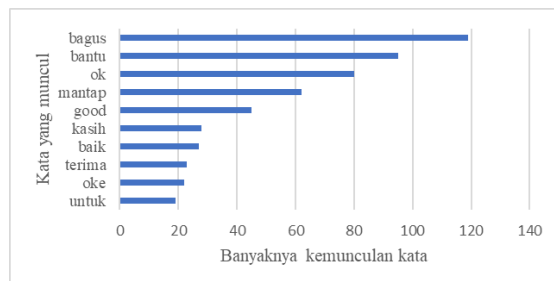
Hasil klasifikasi sentimen ulasan pengguna aplikasi PeduliLindungi pada data uji menunjukkan bahwa dari 1032 kalimat ulasan, terdapat 63.8% (659 ulasan) sentimen positif dan 36.2% (373 ulasan) sentimen negatif. Selanjutnya dari hasil klasifikasi sentimen ulasan, dilihat kata yang paling banyak muncul di setiap kelasnya untuk melihat hal yang paling sering disampaikan oleh pengguna berdasarkan sentimen ulasan.

Tabel 5. Hasil klasifikasi ulasan pengguna pada data uji

No	Ulasan Data Uji	Sentimen Aktual	Sentimen Prediksi
4130	malas banget tidak nyaman loadingnya lama kembali beranda langsung menyahkan	Negatif	Negatif
4131	baik	Positif	Positif
4132	pop up ganggu	Negatif	Positif
⋮	⋮	⋮	⋮
5160	bermanfaat	Positif	Positif
5161	bermanfaat	Positif	Positif
5162	respon whatsapp lama banget	Negatif	Negatif

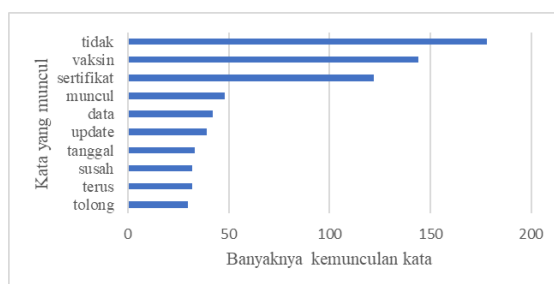
Kata yang paling sering muncul pada kalimat ulasan dengan sentimen positif adalah "bagus", "bantu", "ok", "mantap", dan "good" seperti termuat dalam Gambar 2. Artinya pengguna yang memberikan ulasan positif merasakan kepuasan kepada aplikasi PeduliLindungi dan

pemakaian aplikasi dapat membantu pengguna pada masa pandemi COVID-19. Secara keseluruhan, ulasan positif tidak terlalu mengacu pada aspek khusus dari fitur yang ada di aplikasi PeduliLindungi.



Gambar 2. Kata yang paling banyak muncul pada ulasan positif

Kata yang paling sering muncul pada kalimat ulasan dengan sentimen negatif adalah "tidak", "vaksin", "sertifikat", "muncul", "data", "update", dan "tanggal", seperti termuat dalam Gambar 3. Tidak seperti ulasan positif, pada ulasan negatif pengguna merujuk pada fitur tertentu yang menjadi permasalahan. Fitur yang sering dikeluhkan pengguna adalah fitur sertifikat vaksin, pembaruan (*update*) aplikasi yang dirasa pengguna dilakukan terlalu sering, dan fitur memasukan data tanggal yang dianggap terlalu rumit.



Gambar 3. Kata yang paling banyak muncul pada ulasan negatif

Secara keseluruhan berdasarkan hasil klasifikasi, pemerintah sudah baik dalam mengembangkan aplikasi PeduliLindungi dibuktikan dengan ulasan pengguna yang mengarah kepada sentimen positif. Berdasarkan kata yang paling banyak muncul dalam ulasan sentimen negatif, pada pengembangan selanjutnya direkomendasikan untuk memperbaiki fitur sertifikat vaksin, mengurangi

pembaruan (*update*) aplikasi yang dirasa pengguna dilakukan terlalu sering, dan memperbaiki fitur memasukan data tanggal yang dianggap terlalu rumit.

4.8 Evaluasi Model Klasifikasi

Dari hasil *confusion matrix* yang termuat pada Tabel 6, dihitung nilai akurasi, presisi, *recall*, dan *F1-Score* menggunakan persamaan (11)-(14). Diperoleh nilai akurasi sebesar 93.6%, presisi 93.9%, *recall* 95.9%, dan *F1-Score* 94.9%.

Tabel 6. *Confusion matrix* hasil klasifikasi pada data uji

	Aktual Positif	Aktual Negatif
Prediksi Positif	347	26
Prediksi Negatif	39	620

Untuk melihat rata-rata performa metode klasifikasi model yang diterapkan pada analisis sentimen jika diuji dengan data latih dan data uji yang berbeda, dilakukan evaluasi model menggunakan *10-fold cross validation*. Menggunakan bahasa pemrograman Python, diperoleh nilai metrik dari *10-fold cross validation* dapat dilihat pada Tabel 7.

Tabel 7. Hasil evaluasi menggunakan *10-fold cross validation*

Iterasi	Akurasi	Presisi	Recall	F1-Score
1	94.1%	97.2%	95.9%	96.5%
2	93.2%	94.9%	96.1%	95.5%
3	93.6%	97.0%	94.3%	95.6%
4	94.4%	95.9%	95.0%	95.5%
5	92.4%	95.0%	93.6%	94.3%
6	90.9%	93.0%	92.1%	92.8%
7	94.6%	95.3%	95.9%	95.6%
8	92.6%	92.8%	96.1%	94.4%
9	92.8%	93.9%	95.5%	94.7%
10	95.3%	95.1%	97.0%	96.1%

Hasil akurasi, presisi, *recall*, dan *f1-score* yang dihasilkan sesuai pada Tabel 7 berkisar antara 90%-96%. Karena distribusi kelas data

tidak seimbang (*imbalanced data*), perhitungan evaluasi model lebih cocok dihitung dengan menggunakan metrik *F1-Score*. Berdasarkan Tabel 7, diperoleh nilai rata-rata *F1-Score* model menggunakan *10-fold cross validation* adalah 95.1%.

5. KESIMPULAN

Berdasarkan pembahasan yang telah dilakukan, dengan demikian diperoleh simpulan bahwa hasil analisis sentimen masyarakat menggunakan metode klasifikasi Naïve Bayes dan seleksi fitur *Information Gain* menunjukkan bahwa masyarakat cenderung memberikan ulasan positif terhadap aplikasi PeduliLindungi. Hasil evaluasi performa model untuk analisis sentimen masyarakat terhadap aplikasi PeduliLindungi menggunakan metode klasifikasi Naïve Bayes dan seleksi fitur *Information Gain* berdasarkan rata-rata *F1-Score* dengan *10-fold cross validation* diperoleh sebesar 95.1%.

6. REFERENSI

- Aggarwal, C.C. 2014. *Data Classification: Algorithms and Applications*, *Data Classification: Algorithms and Applications*. New York: CRC Press.
- Aliyah, N. *et al.* 2018. "Colloquial Indonesian Lexicon," in *2018 International Conference on Asian Language Processing (IALP)*. IEEE, hal. 226–229.
- Bird, S., Klein, E. dan Loper, E. 2009. *Natural language processing with Python: Analyzing Text with the Natural Language Toolkit*. 1 ed. Sebastopol: O'Reilly Media, Inc.
- Fauziyyah, N. dan Ersyafdi, I.R. 2021. "Dampak Covid-19 pada Pasar Saham di Berbagai Negara," *FORUM EKONOMI: Jurnal Ekonomi, Manajemen dan Akuntansi*, 23(1), hal. 56–66.
- Firmahsyah, F. dan Gantini, T. 2016. "Penerapan Metode Content-Based Filtering Pada Sistem Rekomendasi Kegiatan Ekstrakurikuler (Studi Kasus di Sekolah ABC)," *Jurnal Teknik Informatika dan Sistem Informasi*, 2(3).
- Haerani, E. dan Rahmatulloh, A. 2021. "Analisis User Experience Aplikasi Peduli Lindungi untuk Menunjang Proses Bisnis Berkelanjutan," *Jurnal SATIN (Sains dan Teknologi Informasi)*, 7(2), hal. 01–10.
- Iqbal, M. *et al.* 2020. "Review of Feature Selection Methods for Text Classification," *International Journal of Advanced Computer Research*, 10(49), hal. 138–152.
- Jović, A., Brkić, K. dan Bogunović, N. 2015. "A Review of Feature Selection Methods with Applications," in *38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. IEEE.
- Jurafsky, D. dan Martin, J. 2019. "Naive Bayes and Sentiment Classification," in *Speech and Language Processing*. New Jersey: Prentice Hall, hal. 1024.
- KOMINFO. 2021. *Aplikasi PeduliLindungi Optimalkan Pengendalian Pandemi*. Tersedia pada: <https://covid19.go.id/p/berita/aplikasi-pedulilindungi-optimalkan-pengendalian-pandemi> (Diakses: 3 Oktober 2021).
- Kurniawati *et al.* 2020. "Public Acceptance of Pedulilindungi Application in the Acceleration of Corona Virus (Covid-19) Handling," in *Journal of Physics: Conference Series*.
- Larose, D.T. 2006. *Data Mining Methods and Models, Data Mining Methods and Models*. New Jersey: John Wiley & Sons.
- Liu, B. 2012. *Sentiment Analysis and Opinion Mining, Synthesis Lectures on Human Language Technologies*. USA: Morgan & Claypool Publishers..
- Liu, B. *et al.* 2013. "Scalable Sentiment Classification for Big Data analysis using Naive Bayes Classifier," in *2013 IEEE International Conference on Big Data*. Santa Clara: IEEE, hal. 99–104.
- Novendri, R. *et al.* 2020. "Sentiment Analysis of

YouTube Movie Trailer Comments Using Naïve Bayes,” *Bulletin of Computer Science and Electrical Engineering*, 1(1), hal. 26–32.

Nugraha, S. dan Chandra, A. 2017. *Stopwords Bahasa Indonesia*. Tersedia pada: github stopwords-bahasa-indonesia (Diakses: 1 Oktober 2021).

Pitria, P. 2014. *Analisis Sentimen Pengguna Twitter Pada Akun Resmi Samsung Indonesia Dengan Menggunakan Naive Bayes*. Universitas Komputer Indonesia.

Tharwat, A. 2020. “Classification Assessment Methods,” *Applied Computing and Informatics*, 17(1).

Walpole, R.E. *et al.* 2016. *Probability & Statistics for Engineers Scientists*. 9 ed, *Education*. 9 ed. Boston: Pearson.