

IMPLEMENTASI METODE *NAIVE BAYES CLASSIFIER* UNTUK KLASIFIKASI SENTIMEN ULASAN PENGGUNA APLIKASI NETFLIX PADA GOOGLE PLAY

Jessica Athalia Rieuwpassa¹, Sugito², Tatik Widiharah³

^{1,2,3}Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro

*e-mail: jessica.athalia07@gmail.com

DOI: 10.14710/j.gauss.12.3.362-371

Article Info:

Received: 2022-09-05

Accepted: 2022-12-02

Available Online: 2024-02-26

Keywords:

Sentiment Analysis; Naive Bayes Classifier; Streaming Service Application; Netflix

Abstract: The COVID-19 pandemic has led to restrictions on activities in public places or facilities, such as cinemas. This has resulted in increased users of streaming service applications such as Netflix where users can access videos or movies online. Netflix users continue to increase from year to year, but its users began to decrease along with other streaming applications. Related to this, sentiment analysis was carried out on the classification of positive and negative reviews given by users on the Google Play website. The classification is expected to produce good accuracy and be analyzed so that it can be useful information for Netflix and potential users of streaming applications. The Naive Bayes Classifier method is a classification algorithm that is easy to apply and has high effectiveness for classifying text. This method utilizes the concept of conditional probability and has a strong assumption of independence. This study uses 2.850 Netflix application review data on Google Play which is then processed and divided into training data and test data with a ratio of 80:20. Classification with the Naive Bayes Classifier produces an accuracy value of 87,33%, a precision value of 87,6%, a recall value of 93,53%, and an F-measure value of 90,47% so it can be concluded that the performance of the Naive Bayes method is good for classifying user reviews of the Netflix.

1. PENDAHULUAN

Layanan *streaming* film dikenal juga dengan *Video on Demand (VoD)* yaitu sistem televisi interaktif yang memberikan kontrol penuh kepada penggunanya untuk dapat memilih program atau video yang diminati untuk ditonton. Salah satu aplikasi layanan *streaming* yaitu Netflix yang didirikan pada tahun 1997 di California dan mulai masuk ke Indonesia pada tahun 2016 (Panji, 2016). Selama 2020 Netflix sukses besar akibat adanya pandemi, namun memasuki 2021 pertumbuhan jumlah pelanggan Netflix melambat. Pihak Netflix mengatakan penurunan jumlah pelanggan baru ini disebabkan pandemi yang menghambat proses pembuatan film-film baru sehingga menurunkan minat pengguna untuk berlangganan (Riyanto, 2021).

Google Play merupakan situs dari perusahaan Google berupa toko digital yang berisi produk-produk aplikasi hiburan, buku, musik, dan film. Google Play ini difasilitasi oleh berbagai fitur-fitur yang dapat kita akses, salah satunya yaitu ulasan aplikasi dari pengguna. Fitur ini dapat bermanfaat untuk mengetahui bagaimana pendapat pengguna yang telah menggunakan suatu aplikasi seperti ulasan pengguna aplikasi Netflix pada situs Google Play. Berdasarkan hal tersebut maka dapat dilakukan analisis sentimen untuk mengetahui pendapat pengguna berdasarkan sentimen ulasan di situs Google Play pada aplikasi Netflix.

Analisis sentimen adalah suatu proses dimana data tekstual dapat dipahami, diekstrak, serta diolah secara otomatis sehingga dapat diperoleh informasi sentimen yang terkandung dalam suatu kalimat opini (Rozi *et al.*, 2012). Banyak hal yang perlu dipersiapkan dahulu untuk membuat analisis, salah satunya yaitu memilih *classifier* yang

akan digunakan. Dalam penelitian ini *classifier* yang dipilih yaitu *Naive Bayes Classifier*. Algoritma *Naive Bayes Classifier* bekerja dengan menghitung probabilitas kelas dan probabilitas bersyarat menggunakan data *training* untuk kemudian digunakan dalam mengklasifikasikan pengamatan baru. Kelebihan metode klasifikasi ini adalah kecepatan waktu proses, kemudahan implementasi dengan strukturnya yang sederhana, serta efektivitasnya tinggi dikarenakan kinerja *Naive Bayes* sering memberikan performa sangat baik saat diaplikasikan dalam permasalahan nyata dan sangat berguna untuk data berdimensi tinggi karena probabilitas setiap fiturnya diasumsikan independen (Taheri dan Mammadov, 2013).

Penulis akan melakukan analisis sentimen dan mengukur performa implementasi metode *Naive Bayes Classifier* terhadap ulasan aplikasi Netflix pada Google Play. Melalui penelitian ini diharapkan dapat mengklasifikasi teks ulasan dengan akurasi yang baik dan dapat mengidentifikasi topik yang sering muncul dalam ulasan sehingga dapat disajikan menjadi informasi yang bermanfaat bagi pihak-pihak yang membutuhkan seperti perusahaan dalam bidang layanan *streaming* film maupun para calon pengguna aplikasi *streaming* film.

2. TINJAUAN PUSTAKA

Proses analisis sentimen diawali dengan *data mining* yang dapat disebut dengan “*knowledge discovery*” ataupun menemukan pola tersembunyi pada data. Bidang *data mining* sudah diaplikasikan pada data berbentuk teks yang dikenal dengan *text mining* yaitu teknologi yang mampu menganalisis data teks semi terstruktur maupun tidak terstruktur (Han *et al.*, 2012). Sebagian besar informasi disimpan sebagai teks seperti artikel berita, makalah teknis, buku, perpustakaan digital, pesan email, blog, dan halaman *web*, sehingga Tujuan pentingnya yaitu untuk memperoleh informasi yang berkualitas dari data berbentuk teks.

Teks yang diproses dengan *text mining* umumnya memiliki beberapa karakteristik seperti terdapat *noise* pada data dan terdapat struktur teks yang kurang baik, untuk memperoleh bentuk data yang siap untuk diolah melalui *data mining*, maka terlebih dahulu perlu dilakukannya teks *preprocessing*. Terdapat beberapa tahapan yang dilakukan untuk *preprocessing* data diantaranya *spelling normalization*, *case folding*, filtrasi, tokenisasi, dan *stemming*.

1. *Spelling Normalization*

Spelling Normalization merupakan proses memperbaiki ejaan yang tidak tepat ataupun kata-kata yang berbentuk singkatan tertentu. Contoh dari kata-kata yang perlu diperbaiki yaitu seperti “yg”, “gak”, “bgus” dan lain sebagainya.

2. *Case Folding*

Case folding merupakan proses untuk mengubah seluruh kata menjadi bentuk yang sama. Pada tahapan ini, seluruh huruf akan diubah menjadi huruf kecil semua (Hidayatullah & Ma'Arif, 2017).

3. Filtrasi

Tahap filtrasi yaitu proses penghapusan tanda yang tidak memiliki makna atau arti secara signifikan untuk pengolahan data selanjutnya, seperti tanda baca, angka, dan simbol. Selanjutnya dilakukan *stopword removal* yaitu menghilangkan kata-kata umum yang biasanya tidak memiliki pengaruh signifikan dalam kalimat menggunakan *stopwords* Bahasa Indonesia contohnya seperti “dan”, “atau”, dan sebagainya (Hidayatullah dan Ma'Arif, 2017).

4. Tokenisasi

Tokenisasi bertujuan untuk membagi kalimat menjadi beberapa bagian yang disebut dengan token. Token ini dapat berupa kata, frasa ataupun unsur makna lainnya (Hidayatullah dan

Ma'Arif, 2017). Hasil dari proses ini yaitu menjadikan kalimat-kalimat pada dokumen teks menjadi kumpulan kata yang terpisah.

5. Stemming

Stemming adalah proses yang bertujuan untuk memperoleh kata dasar atau akar kata dengan cara menghilangkan imbuhan dan sufiks pada kata-kata dalam teks (Hidayatullah dan Ma'Arif, 2017). Contoh hasil dari *stemming* yaitu “menunggu” menjadi “tunggu”, “pembayaran” menjadi “bayar”, “dibuka” menjadi “buka”, dan sebagainya.

Pelabelan data dilakukan dengan teknik *sentiment scoring* berdasarkan kamus *lexicon* dimana data diberikan label positif maupun negatif berdasarkan total skornya. Fitur *Lexicon-Based* adalah fitur kata yang mempunyai sentimen positif ataupun negatif yang bersumber pada kamus atau *lexicon*. *Wordlist* dalam kamus sentimen tersebut memiliki bobot dengan kekuatan sentimen 1 sampai 5 (memiliki sentimen positif), serta -1 sampai -5 (memiliki sentimen negatif) yang menjadi dasar perhitungan skor suatu dokumen.

Term Frequency Inverse Document Frequency (TF-IDF) adalah metode pembobotan kata yang menggabungkan metode *TF* dan *IDF*. Metode ini merupakan sebuah kombinasi metode yang dapat memberikan performansi yang lebih baik, khususnya dalam memperbaiki nilai *recall* dan *precision* (Deolika et al., 2019). Rumus umum *TF-IDF* diperoleh dengan menggabungkan perhitungan *TF* murni dan *IDF* menjadi persamaan sebagai berikut (Ghag dan Shah, 2014)

$$w_{i,d} = \frac{n_{i,d}}{\sum_k n_{k,d}} \times \log_2 \left(\frac{D}{df_i} \right) \quad (1)$$

dengan D merupakan jumlah seluruh dokumen, $n_{i,d}$ yaitu jumlah *term* ke i muncul pada dokumen ke d , $\sum_k n_{k,d}$ yaitu jumlah seluruh *term* yang terdapat pada dokumen ke d , dan df_i yaitu jumlah dokumen yang mengandung *term* i .

Naive Bayes merupakan bentuk prediksi peluang di masa mendatang berdasarkan pengalaman yang terjadi di masa sebelumnya. Prinsip umum *Naive Bayes Classifier* mengasumsikan bahwa nilai dari suatu atribut tidak bergantung dan mempengaruhi atribut lainnya (Pintoko, 2018). Model *Naive Bayes* yang digunakan yaitu *Multinomial Naive Bayes* dengan persamaan umum pada persamaan (2) (Manning et al., 2009)

$$P(V_j|X_i) = P(V_j) \prod_{i=1}^n P(X_i|V_j) \quad (2)$$

Pengklasifikasian dokumen ke suatu kelas akan mencari probabilitas tertinggi dari semua kelas yang diujikan. Untuk mencari probabilitas terbesar dapat dituliskan seperti persamaan (3) (Manning et al., 2009)

$$V_{MAP} = \underset{V_j \in V}{\operatorname{argmax}} P(V_j) \prod_{i=1}^n P(X_i|V_j) \quad (3)$$

dengan $P(V_j)$ merupakan Probabilitas *prior* dari kelas j dan $P(X_i|V_j)$ merupakan Probabilitas kata X_i untuk kelas j (*conditional probability*).

Probabilitas *prior* (probabilitas kelas sentimen) ditentukan dengan rumus persamaan (4) (Manning et al., 2009)

$$P(V_j) = \frac{N_j}{N} \quad (4)$$

Keterangan :

- V_j : kelas klasifikasi
j = 0 (kelas negatif); j = 1 (kelas positif)
- N_j : Jumlah dokumen pada kelas j
- N : Jumlah dokumen dari semua kelas

Probabilitas kemunculan tiap *term* untuk tiap kelas klasifikasi (conditional probability) ditentukan dengan teknik *laplace smoothing* untuk mencegah nilai nol dalam perhitungan probabilitas (Manning *et al.*, 2009)

$$P(X_i|V_j) = \frac{w_{ij} + 1}{(\sum_{w' \in T} w_{ij}') + B} \quad (6)$$

Keterangan:

- w_{ij} : bobot TF-IDF *term* i pada dokumen kelas j
- $\sum_{w' \in T} w_{ij}'$: jumlah bobot TF-IDF *term* pada dokumen kelas j
- B : jumlah IDF seluruh *term* pada dokumen

Proses evaluasi dilakukan guna melihat efektifitas dan kualitas dari model yang digunakan (Ginantra *et al.*, 2021). Metode perhitungan yang bisa digunakan untuk mengevaluasi model klasifikasi yaitu *Confusion Matrix*. Pada evaluasi dengan *confusion matrix* terdapat empat kemungkinan yang dapat muncul dari hasil klasifikasi suatu data yaitu (1) *True Positive (TP)* jika data positif dan diprediksi positif, (2) *False Negative (FN)* jika data positif diprediksi negatif, (3) *True Negative (TN)* jika data negatif diprediksi negatif, dan (4) *False Positive (FP)* jika data negatif diprediksi positif (Sari, 2019).

Tabel 1. *Confusion Matrix*

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

Nilai yang ada pada Tabel 1 kemudian dapat dijadikan acuan untuk menghitung *accuracy*, *precision*, *recall*, serta *f-measure*.

a. *Accuracy*

Accuracy adalah tingkat kedekatan antara nilai prediksi dengan nilai aktual.

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

b. *Precision*

Precision adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban sistem.

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

c. *Recall*

Recall bertujuan untuk melihat ukuran tingkat keberhasilan sistem klasifikasi dalam menemukan kembali sebuah informasi.

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

d. *F-Measure*

F-measure merupakan rata-rata harmonis dari *precision* dan *recall*.

$$F - measure = \frac{2 (recall . precision)}{(recall + precision)}$$

Word Cloud (awan kata) adalah visualisasi data teks yang merupakan sekumpulan kata dengan frekuensi paling banyak muncul dalam teks dokumen yang dianalisis. Dalam *word cloud* pula terlihat perbedaan ukuran huruf dari kata-kata di dalamnya, semakin besar ukuran hurufnya maka menunjukkan semakin sering kata itu muncul pada teks dokumen. Visualisasi menggunakan *word cloud* ini lebih menarik untuk dilihat dan memudahkan untuk menemukan kata-kata yang sering muncul, namun kekurangannya yaitu tidak ditunjukkan frekuensi kemunculan kata-kata dalam teks yang dianalisis (Adiyana & Hakim, 2015).

3. METODOLOGI PENELITIAN

Jenis data yang digunakan dalam penelitian ini yaitu data sekunder yang diperoleh dengan menggunakan teknik *scraping* dengan *Python* melalui *Jupyter Notebook*. Sumber data dari situs *website* Google Play dengan alamat *website* yaitu <https://play.google.com/store/apps/details?id=com.netflix.mediaclient&hl=in&gl=US>. Data yang digunakan berupa ulasan atau *review* pengguna aplikasi Netflix yang diambil sejak 13 Desember 2021 – 7 Maret 2022.

Variabel dalam penelitian ini adalah *Date* (Waktu) yaitu tanggal ulasan diunggah, *Review* (Ulasan) yaitu berisi pendapat dari pengguna mengenai aplikasi Netflix, dan *Rating* (Peringkat) yaitu penilaian pengguna terhadap aplikasi Netflix dalam bentuk angka dengan rentang penilaian bintang 1 (sangat bagus) hingga penilaian bintang 5 (sangat buruk).

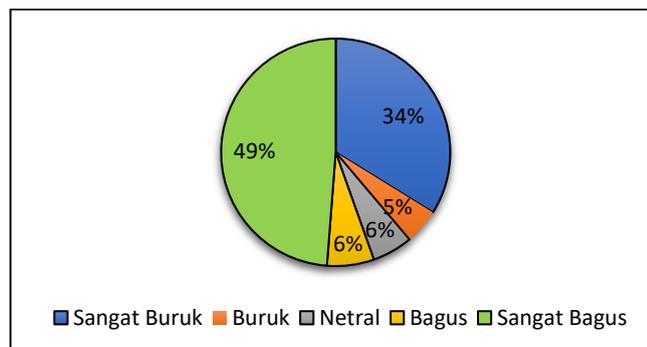
Pengolahan data pada penelitian ini menggunakan *Python* melalui *Jupyter Notebook* dan *Microsoft Excel*. Data yang diperoleh akan dianalisis dengan tahapan sebagai berikut:

1. Mengumpulkan data dengan *scraping* data ulasan aplikasi Netflix dari *website* Google Play
2. Membuat gambaran umum persepsi pengguna terhadap aplikasi Netflix
3. *Preprocessing* data teks dengan lima tahapan yaitu *Spelling Normalization*, *Case Folding*, *Filtrasi*, *Tokenisasi*, dan *Stemming*
4. Pelabelan teks ulasan menjadi dua kelas sentimen yaitu positif dan negatif dengan *sentiment scoring*, dengan ketentuan jika skor > 0 yaitu positif, skor $= 0$ yaitu netral, dan skor < 0 yaitu negatif
5. Pembobotan kata menggunakan *Term Frequency-Inverse Document Frequency (TF-IDF)*.
6. *Split* data menjadi data latih dan data uji dengan ratio 80:20
7. Membuat klasifikasi dengan algoritma *Naive Bayes Classifier*
8. Evaluasi hasil klasifikasi dengan *confusion matrix* lalu menghitung nilai *accuracy*, *precision*, *recall*, dan *F-measure*
9. Visualisasi hasil klasifikasi dalam bentuk *word cloud*

4. HASIL DAN PEMBAHASAN

Pengumpulan data ulasan pengguna aplikasi Netflix pada *website* Google Play dilakukan dengan teknik *web scraping* yaitu menggunakan *google play scraper* pada *Python*. Data ulasan yang diambil yaitu 12 minggu atau 3 bulan (13 Desember 2021 - 7 Maret 2022), kemudian diperoleh 2.850 data ulasan dengan 3 variabel yaitu '*content*' dimana berisi ulasan, '*at*' berisi tanggal ulasan diunggah, dan '*score*' berisi rating yang diberikan pengguna. Hasil *web scraping* dari *Python* kemudian disimpan dalam format *CSV (Comma Separated Value)*.

Berikut gambaran umum dari data ulasan aplikasi Netflix menurut *rating* yang diberikan pengguna



Gambar 1. Rating Aplikasi Netflix Desember 2021 – Maret 2022

Berdasarkan grafik *pie chart* pada Gambar 1 dari total 2.850 ulasan pengguna untuk aplikasi Netflix paling banyak memberikan bintang 5 (sangat bagus) dengan persentase 49% atau sebanyak 1.390 pengguna, diikuti persentase terbesar kedua yaitu bintang 1 (sangat buruk) dengan persentase 34% atau sebanyak 965 pengguna yang artinya disamping banyak pengguna

yang memberikan ulasan positif, banyak pula pengguna yang memberikan ulasan negatif yang biasanya berisi kritikan atau keluhan kurang puas terhadap aplikasi Netflix.

Tahap *preprocessing* teks dilakukan beberapa tahapan agar data siap untuk diolah ke proses selanjutnya. *Preprocessing* teks terdiri dari 5 tahapan yaitu *spelling normalization*, *case folding*, filtrasi, tokenisasi, dan *stemming*. Berikut ini hasil *output preprocessing* teks ulasan pada *Python*.

	ulasan	tokenization	ulasan_stem	word_length
0	aplikasi film mengesankan digunakan sempurna	[aplikasi, film, mengesankan, digunakan, sempu...	aplikasi film kesan guna sempurna	5
1	astaga si ganti pembayaran bayar rb buka nge...	[astaga, si, ganti, pembayaran, bayar, rb, buk...	astaga si ganti bayar bayar rb buka ngeselin r...	17
2		[ly]	ly	1
3	mantap	[mantap]	mantap	1
4	aplikasi kontrol	[aplikasi, kontrol]	aplikasi kontrol	2
...
2845	fu setingan subtitle warna nyaru video ditampl...	[fu, setingan, subtitle, warna, nyaru, video, ...	fu ting subtitle warna nyaru video tampli	7
2846	bagus	[bagus]	bagus	1
2847	hapus	[hapus]	hapus	1
2848	pengalaman coba	[pengalaman, coba]	alam coba	2
2849	bagus	[bagus]	bagus	1

2850 rows x 4 columns

Gambar 2. Output Hasil *Preprocessing* Teks

Hasil *preprocessing* data pada Gambar 2 yaitu variabel “ulasan_stem” yang akan digunakan untuk pelabelan. Penilaian skor dokumen dilakukan berdasarkan *wordlist* dalam kamus *lexicon* yang sudah diberikan bobot dengan cara menjumlahkan nilai bobot seluruh kata pada dokumen teks ulasan, kemudian menghasilkan tiga kelas sentimen yaitu sentimen positif, sentimen netral, dan sentimen negatif. Pada penelitian ini hanya akan digunakan dua kelas yaitu positif dan negatif sehingga kelas netral akan direduksi dengan beberapa pertimbangan. Berikut hasil pelabelan kelas sentimen setelah hasil reduksi kelas sentimen netral ke kelas positif dan negatif

Tabel 2. Hasil Pelabelan Kelas Sentimen

Kelas Sentimen	Jumlah Ulasan
Positif	1.631
Negatif	1.010
Jumlah	2.641

Berdasarkan Tabel 2 di atas diperoleh bahwa dari total 2.641 ulasan, yang terdiri dari 1.631 ulasan positif, dan 1.010 ulasan negatif.

Perhitungan pembobotan kata (*term*) dengan menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)* untuk seluruh kata pada data ulasan aplikasi Netflix secara otomatis dengan bantuan *Python* menggunakan *syntax TfidfVectorizer* dari *sklearn.feature_extraction*, kemudian dihasilkan *output* pembobotan *TF-IDF* seperti berikut

Tabel 3. Hasil Pembobotan TF-IDF

Dokumen	aplikasi	bagus	bayar	film	mantap	zaman
D1	0,202706	0	0	0,251317	0	...	0
D2	0	0	0,241232	0	0	...	0
D3	0	0	0	0	1	...	0
D4	0,386736	0	0	0	0	...	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
D2641	0	1	0	0	0	...	0

Berdasarkan *output* pada Tabel 3 diperoleh nilai bobot untuk 2.502 kata (*term*) pada setiap dokumen dari dokumen 1 (d1) hingga dokumen 2641 (d2641).

Setelah kata-kata dalam teks ulasan diberikan nilai dan bobot untuk setiap kata kemudian data dibagi menjadi data latih dan data uji. Berdasarkan *Pareto Principle*, perbandingan yang paling umum digunakan yaitu 80:20 dalam membagi data latih dan data uji

(Dunford *et al.*, 2014) sehingga pada penelitian ini perbandingan jumlah data latih dan data uji yang digunakan yaitu 80:20. Pembagian data dilakukan dengan *Python* menggunakan *train_test_split* dari *sklearn.model_selection* dan menghasilkan pembagian data sebagai berikut

Tabel 4. Pembagian Data Latih dan Data Uji

Klasifikasi	Jumlah	Data Latih (80%)	Data Uji (20%)
Positif	1.631	1.291	340
Negatif	1.010	821	189
Total	2.641	2.112	529

Berdasarkan Tabel 4 dengan perbandingan data latih dan data uji sebesar 80:20 dari total 2.641 data ulasan yang digunakan yaitu sebanyak 2.112 data sebagai data latih yang terdiri dari 1.291 data ulasan positif dan 821 data ulasan negatif, kemudian 529 data sebagai data uji yang terdiri dari 340 data ulasan positif dan 189 data ulasan negatif.

Klasifikasi data ulasan dengan *MultinomialNB* dari *sklearn.naive_bayes* pada *Python*. Klasifikasi kelas sentimen ini menghasilkan *output* prediksi kelas sentimen seperti berikut:

Tabel 5. Hasil Prediksi *Naive Bayes Classifier*

No.	Index Ulasan	Prediksi
0	2338	0
1	32	1
2	1681	1
3	1462	1
4	478	1
...
524	279	1
525	2414	0
526	2357	1
527	1874	1
528	24	0

Pada Tabel 5 di atas menunjukkan *output* hasil prediksi kelas sentimen untuk data uji dengan klasifikasi nilai “0” adalah ulasan negatif dan nilai “1” adalah ulasan positif. Setelah diperoleh prediksi klasifikasi dari algoritma *Naive bayes*, selanjutnya akan dilakukan evaluasi untuk mengukur kinerja metode *Naive Bayes* dalam mengklasifikasikan data ulasan aplikasi Netflix. Berikut hasil *confusion matrix* yang diperoleh:

Tabel 6. Hasil *Confusion Matrix*

Aktual	Prediksi		Total
	Negatif	Positif	
Negatif	144	45	189
Positif	22	318	340
Total	166	363	529

$$(1) \text{ Accuracy} = \frac{TP+TN}{(TP+TN+FP+FN)} \times 100\% = \frac{318+144}{(318+144+45+22)} \times 100\% = 87,33\%$$

$$(2) \text{ Precision} = \frac{TP}{(TP+FP)} \times 100\% = \frac{318}{(318+45)} \times 100\% = 87,6\%$$

$$(3) \text{ Recall} = \frac{TP}{(TP+FN)} \times 100\% = \frac{318}{(318+22)} \times 100\% = 93,53\%$$

$$(4) \text{ F-Measure} = \frac{2(\text{recall} \cdot \text{precision})}{(\text{recall}+\text{precision})} = \frac{2(93,53 \cdot 87,6)}{(93,53+87,6)} = 90,47\%$$

Hasil evaluasi kinerja metode *Naive Bayes Classifier* diperoleh akurasi sebesar 87,33%, artinya 87,33% data dapat terklasifikasikan dengan benar oleh sistem. Nilai *precision* sebesar 87,6% yang menunjukkan prediksi dari sistem klasifikasi yang tepat sasaran. Nilai *recall* diperoleh sebesar 93,53% berarti keberhasilan sistem untuk dapat mengenali kembali sampel positif pada data uji. Selain itu, sebagai penilaian yang lebih seimbang dari nilai *precision* dan *recall* diperoleh nilai *F-measure* sebesar 90,47%. Nilai *accuracy*, *precision*, *recall*, dan *F-measure* seluruhnya menunjukkan lebih dari 80%, maka dapat disimpulkan bahwa metode *Naive Bayes Classifier* ini memiliki performa kinerja yang baik dan dapat diterapkan untuk klasifikasi sentimen data teks ulasan aplikasi Netflix pada penelitian ini.



Gambar 3. *Word Cloud* Kelas Sentimen Positif

Gambar 3 merupakan visualisasi data menggunakan word cloud untuk kelas positif terlihat bahwa kata yang berukuran lebih besar dibanding kata lainnya yaitu ada kata “bagus”, “film”, “bayar”, “mantap”, dan “banget” artinya kata-kata tersebut sering muncul dalam ulasan positif. Berikut tabel frekuensi kata yang berisi lima kata dengan kemunculan terbanyak dalam sentimen positif:

Tabel 7. Frekuensi Kata Terbanyak Kelas Sentimen Positif

Kata	Frekuensi
bagus	646
film	219
bayar	199
mantap	125
banget	115

Berdasarkan Tabel 7 dapat dilihat kata “bagus” yang menjadi kata yang paling banyak digunakan dan juga terdapat kata “mantap” menunjukkan bahwa pada sentimen positif pengguna beropini bahwa aplikasi Netflix adalah aplikasi yang bagus dan mantap. Selain itu juga, kata yang banyak muncul yaitu “film”, diikuti oleh kata “bayar” yang sebagian besar berasal dari kata “berbayar”. Kata “banget” sering muncul bersamaan dengan kata “bagus” sehingga memiliki frekuensi kemunculan yang banyak juga. Hal ini juga menunjukkan pada sentimen positif pengguna beropini mengenai aplikasi Netflix memiliki koleksi film yang bagus dan aplikasi Netflix merupakan aplikasi untuk menonton film yang bagus meskipun harus membayar. Berdasarkan ulasan-ulasan positif tersebut dapat disimpulkan bahwa harapannya Netflix dapat terus menjaga dan meningkatkan kualitas layanan aplikasinya terutama untuk film yang disediakan agar terus *update* dan bertambah koleksi film yang tersedia.



Gambar 4. *Word Cloud* Kelas Sentimen Negatif

Pada Gambar 4 merupakan visualisasi data menggunakan word cloud untuk kelas negatif terlihat kata yang berukuran lebih besar dibanding kata lainnya yaitu ada kata “aplikasi”, “film”, “bayar”, “ribet”, dan “masuk” artinya kata-kata tersebut sering muncul dalam ulasan negatif. Berikut tabel frekuensi kata yang berisi lima kata dengan kemunculan terbanyak dalam sentimen negatif:

Tabel 8. Frekuensi Kata Terbanyak Kelas Sentimen Positif

Kata	Frekuensi
aplikasi	456
masuk	217
bayar	153
ribet	126
film	113

Berdasarkan Tabel 8 dapat dilihat kata “aplikasi” yang paling banyak muncul merupakan objek dalam ulasan ini yaitu Aplikasi Netflix. Kata lain yang terbanyak yaitu terdapat kata “masuk”, lalu kata “bayar” yang juga sebagian berasal dari kata “pembayaran”, dan kata “ribet” menunjukkan bahwa pada sentimen negatif ini pengguna beropini bahwa untuk dapat masuk ke aplikasi Netflix dan menikmati layanan diharuskan untuk berlangganan dengan membayar terlebih dahulu, sistem atau metode pembayarannya ribet, dan banyak pengguna kesulitan untuk masuk atau mendaftar ke aplikasi. Selain itu, kata yang banyak muncul lagi yaitu kata “film”, akibat sulitnya masuk ke aplikasi Netflix sehingga sulit juga untuk ingin mengakses film dan menikmati layanan dari aplikasi Netflix. Harapannya Netflix dapat lebih memudahkan lagi untuk sistem pembayarannya seperti menambah metode pembayaran dan dapat memperbaiki lagi sistem untuk masuk (*login*) pada aplikasinya.

5. KESIMPULAN

Penerapan klasifikasi sentimen dilakukan untuk data ulasan aplikasi Netflix dengan metode *Naive Bayes Classifier* menggunakan 2.112 data latih dan 529 data uji dengan ratio 80:20. Berdasarkan hasil prediksi klasifikasi diperoleh tingkat akurasi sebesar 87,33% atau dari 529 data ulasan sebanyak 462 data terklasifikasikan dengan benar. Hasil prediksi klasifikasi juga menghasilkan nilai *precision* sebesar 87,6%, nilai *recall* sebesar 93,53%, dan nilai *F-measure* sebesar 90,47%. Berdasarkan nilai *accuracy*, *precision*, dan *recall* ini menunjukkan bahwa performa kinerja metode *Naive Bayes Classifier* ini baik dan dapat diterapkan untuk mengklasifikasikan data ulasan aplikasi Netflix.

Informasi yang diperoleh dari hasil klasifikasi sentimen dan visualisasi data teks ulasan yaitu dapat diketahui bahwa pengguna aplikasi Netflix dalam ulasan positif banyak membicarakan mengenai “bagus”, “film”, “bayar”, “mantap”, dan “banget”. Pada ulasan negatif banyak pengguna membicarakan mengenai “aplikasi”, “masuk”, “bayar”, “ribet”, dan “film”.

DAFTAR PUSTAKA

- Adiyana, I., & Hakim, F. 2015. Implementasi Text Mining pada Mesin Pencarian Twitter Untuk Menganalisis Topik-Topik Terkait "KPK dan Jokowi". *Prosiding Seminar Nasional Matematika dan Pendidikan Matematika UMS 2015*, 570 - 581.
- Deolika, A., Kusriani, & Luthfi, E. T. 2019. Analisis Pembobotan Kata pada Klasifikasi Text Mining. *Jurnal Teknologi Informasi*, 3(2), 179 - 184.
- Dunford, R., Su, Q., & Tamang, E. 2014. The Pareto Principle. *The Plymouth Student Scientist*, 7(1), 140 - 148.
- Ghag, K., Shah, K. 2014. SentiTFIDF - Sentiment Classification using Relative Term Frequency Inverse Document Frequency. *International Journal of Advanced Computer Science and Applications*, 5(2), 36 - 43.
- Ginantra, N. L. W. S. R., Arifah, F. N., Wijaya, A. H., Septarini, R. S., Ahmad, N., Ardiana, D. P. Y., & Negara, E. S. 2021. *Data Mining dan Penerapan Algoritma*. Yayasan Kita Menulis.
- Han, J., Kamber, M., & Pei, J. 2012. *Data Mining Concepts and Techniques* (3 ed.). USA: Morgan Kaufmann.
- Hidayatullah, A. F., & Ma'arif, M. R. 2017. Pre-processing Tasks in Indonesia Twitter Messages. *Journal of Physics: Conference Series*, 801. doi:10.1088/1742-6596/801/1/012072.
- Manning, C., Raghavan, P., & Schütze, H. 2009. *An Introduction to Information Retrieval*. Cambridge: Cambridge University Press.
- Panji, A. 2016. *Netflix Resmi Tersedia di Indonesia*. Dalam CNN Indonesia: <https://www.cnnindonesia.com/teknologi/20160107083412-185-102757/netflix-resmi-tersedia-di-indonesia> Diakses pada 12 Maret 2022.
- Pintoko, B. M., L., K. M. 2009. Analisis Sentimen Jasa Transportasi Online pada Twitter Menggunakan Metode Naive Bayes Classifier. *e-Proceedings of Engineering*, 5(3), 8121 - 8130.
- Riyanto, G. P. 2021. *Pertumbuhan Pelanggan Netflix Tidak Sesuai Harapan*. Dalam Kompas.com: <https://tekno.kompas.com/read/2021/04/22/12100017/pertumbuhan-pelanggan-netflix-tidak-sesuai-harapan?page=all> Diakses pada 12 Maret 2022.
- Rozi, I. F., Pramono, S. H., & Dahlan, E. A. 2012. Implementasi Opinion Mining (Analisis Sentimen) untuk Ekstraksi Data Opini Publik pada Perguruan Tinggi. *Jurnal EECCIS*, 6(1), 37-43.
- Sari, A. E., Widowati, S., & Lhaksmana, K. M. 2019. Klasifikasi Ulasan Pengguna Aplikasi Mandiri Online di Google Play Store dengan Menggunakan Metode Information Gain dan Naive Bayes Classifier. *eProceedings of Engineering*, 6(2), 9143 - 9157.
- Taheri, S., & Mammadov, M. 2013. Learning The Naive Bayes Classifier with Optimization Models. *International Journal of Applied Mathematics and Computer Science*, 23(4), 787 - 795.