

Naskah Publikasi-Fina Khoirunnisa-18.01.55.0044- 01082022

by Tete Hayati

Submission date: 01-Aug-2022 11:08AM (UTC+0700)

Submission ID: 1877532704

File name: NASKAH_PUBLIKASI_Fina_Khoirunnisa_1801550044.pdf (1.17M)

Word count: 4792

Character count: 29289



ANALISIS SENTIMEN KUALITAS LAYANAN GOOGLE MEET MENGGUNAKAN *NAÏVE BAYES* *CLASSIFIERS* DAN *ASSOCIATION*

Fina Khoirunnisa^{*1}, Herny Februariyanti²

^{1,2}Program Studi Sistem Informasi, Fakultas Teknologi Informasi dan Industri, Universitas Stikubank
e-mail: ^{*1}finakhoirunnisa@mhs.unisbank.ac.id, ²hernyfeb@edu.unisbank.ac.id

Abstrak

Teknologi sangat berperan besar dalam masa pandemi COVID-19 dimana pemerintah mengeluarkan kebijakan pembatasan sosial dalam pengendalian penyebaran virus COVID-19. Pandemi mempengaruhi segala proses aktivitas normal sehingga masyarakat memerlukan *digital video conferencing* untuk menjalin komunikasi visual dalam kelompok besar yang ditandai dengan meningkatnya penggunaan aplikasi sejenis yaitu *Google Meet*. Analisis sentimen merupakan proses untuk mendapat informasi sentimen yang terkandung dalam sebuah kalimat opini. Sampel data yang digunakan adalah ulasan aplikasi *Google Meet* pada situs *Google Play* berdasarkan data ulasan dari tanggal 1 Januari 2021 sampai 25 Agustus 2021. Analisis sentimen dilakukan menggunakan metode *Naïve Bayes Classifiers* dan *text Association*. Didapat persentase sebesar 58,8% dan 41,2% dari sentimen positif dan negatif. Dengan menggunakan perbandingan 80:20 untuk data latih dan uji. Didapatkan hasil akurasi sebesar 85,65% untuk pengujian dengan *Naïve Bayes* dan 85,48% untuk hasil validasi model menggunakan *10-Fold Cross-Validation*. Secara umum hasil dari *text Association* yang dihasilkan merupakan ekstraksi informasi kelas positif terkait *meeting*, video, *problem*, koneksi, suara, audio, kualitas, kamera, dan fitur. Sedangkan untuk kelas negatif menampilkan ekstraksi dari hal yang sering dikeluhkan terkait video, masalah, koneksi, audio, suara, *update*, kualitas, *camera*, dan fitur. Dengan penelitian tersebut sehingga dapat menjadi bahan acuan dalam upaya untuk menjaga dan meningkatkan kualitas aplikasi.

Kata kunci; Analisis Sentimen, Kualitas Layanan, Pengklasifikasi *Naïve Bayes*, Teks Asosiasi

Abstract

Technology plays a big role in the COVID-19 pandemic where the government issued a social restriction policy to control the spread of the COVID-19. Pandemic affects all normal activity processes so people need digital video conferencing to establish visual communication in large groups which is marked by an increase of similar applications namely Google Meet. Sentiment analysis is the process of obtaining sentiment information in an opinion sentence. The data sample is a review of the Google Meet application on the Google Play site-based review data from January,1,2021 to August,25,2021. Sentiment analysis was conducted using Naïve-Bayes-Classifiers and text-Association methods. The percentage of positive and negative sentiment is 58.8% and 41.2%. Using a ratio of 80:20 for training and test data. The results are 85.65% accuracy for tests with Naïve Bayes and 85.48% for model validation using 10-Fold Cross-Validation. In general, the results of text Association are positive class information related to meetings, videos, problems, connections, sound, audio, quality, cameras, and features. The negative class displays things that are often complained about related to video, problems, connections, audio, sound, updates, quality, camera, and features. This research can be used as a reference to maintain and improve the quality of the application.

Keywords; Sentiment Analysis, Service Quality, *Naïve Bayes Classifier*, Association Text



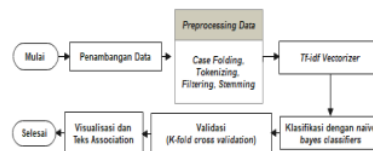
1. PENDAHULUAN

Percepatan era revolusi *industry 4.0* dalam perkembangan teknologi digitalisasi berkembang begitu pesat. Teknologi berperan besar dalam masa pandemi COVID-19 dimana pemerintah mengeluarkan kebijakan kepada masyarakat untuk menerapkan pembatasan sosial guna mengendalikan dan memperlambat penyebaran virus COVID-19. Pandemi telah meningkatkan popularitas *video conference* di kalangan masyarakat umum, terbukti dengan peningkatan jumlah pengguna yang signifikan. *Video conference* merupakan salah satu jenis layanan multimedia yang dapat memenuhi kebutuhan *user* yang menginginkan adanya komunikasi secara *real time, anytime, and anywhere*. Layanan ini dapat diimplementasikan melalui suatu jaringan (*network*) [1]. Aplikasi *Google Meet* adalah salah satu aplikasi *video conference* paling populer dan banyak digunakan oleh beberapa kelompok. Untuk mendapatkan informasi produk, diperlukan *reviews* pengguna sebagai alat efektif untuk mencari informasi. Sebagian besar calon pengguna sangat bergantung pada rekomendasi atau ulasan dari pengguna lain sebelum menggunakan produk yang akan digunakan, karena perspektif dari pengguna lain mempengaruhi dan memberikan informasi baru terhadap suatu produk yang ingin digunakan. Menggali informasi tentang sentimen dapat memberitahu konsumen dan produsen tentang bagaimana pasar bereaksi. Sehingga memungkinkan analisis dan tindakan yang diperlukan berdasarkan informasi yang ditemukan. Mengumpulkan dan memproses semua *reviews* secara manual membutuhkan banyak waktu, sehingga diperlukan cara yang tepat untuk menangani *reviews*. *Text mining* dapat digunakan untuk mengklasifikasikan dokumen teks berupa opini berdasarkan sentimen [2]. Analisis sentimen dapat mempelajari opini atau pendapat yang menyuarakan atau mengungkapkan persepsi yang mengandung sentimen positif atau negatif [3]. Algoritma *Naïve Bayes* merupakan algoritma paling sederhana dari pengklasifikasi probabilistik [4] dan memiliki tingkat kesalahan yang sangat minimum dibanding dengan algoritma lainnya [5]. Dalam penelitian ini, peneliti melakukan analisis

sentimen kualitas layanan terhadap aplikasi *Google Meet* berdasarkan hasil *scraping* yang diperoleh dari data *reviews* dalam situs *Google Play*. Dengan menggunakan algoritma *Naïve Bayes Classifiers* dan *Association* untuk menyajikan sebuah informasi dan mengetahui persepsi pengguna terhadap kualitas layanan aplikasi *Google Meet*. Berdasarkan penelitian terdahulu yang telah dilakukan oleh Oktaviani, Warsito, Yasin, Santoso, dan Suparti [6] dengan melakukan analisa sentimen terhadap aplikasi *Traveloka* menggunakan data *reviews* dari situs *Google Play*. Hasil akurasi tertinggi dari algoritma *Naïve Bayes Classifier* dengan memanfaatkan *10-Fold Cross-Validation* adalah 91,20%. Kemudian hasil dari ekstraksi asosiasi kata pada *positive reviews* yaitu terkait dengan “promosi”, “cepat”, “keren”, dan “tiket” sedangkan pada *negative reviews* terkait dengan “aplikasi”, “bayar”, “sistem”, dan “pesan”.

2. METODE PENELITIAN

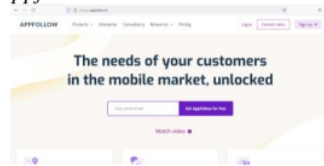
Rencana alur analisis sentimen pada penelitian ini memiliki tahapan alur penelitian yang tersusun secara sistematis sebagaimana alur penelitian ini akan berjalan. Alur tersebut dapat ditunjukkan oleh Gambar 1.



Gambar 1 Diagram Alur Penelitian

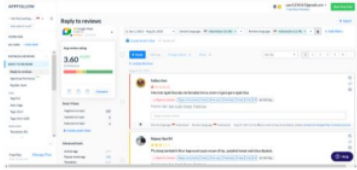
2.1 Penambangan Data

Penambangan data dijalankan dengan melakukan *web scraping* untuk mempermudah serta mengefisiensikan waktu. Pada penelitian ini dilakukan proses *scraping* dengan menggunakan layanan aplikasi berbasis web yaitu *appfollow.io*.



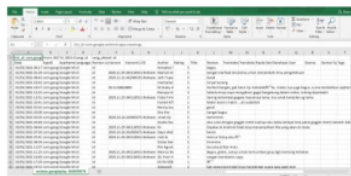
Gambar 2 Halaman Beranda *Appfollow.io*

Ditujukan halaman beranda dari *appfollow.io* pada Gambar 2. Untuk melakukan akses pada situs *Appfollow* dapat dilakukan dengan mengakses alamat <https://appfollow.io/>.



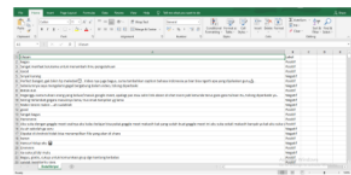
Gambar 3 Proses Pencarian Data Ulasan pada Halaman *Reply to Reviews*

Dalam Gambar 3 ditampilkan proses pencarian data ulasan berdasarkan rentang waktu yang dibutuhkan. selanjutnya data *reviews* dapat diekstrak dengan proses *export* yang nantinya data hasil *scraping* akan disimpan dalam file *.xlsx* atau *.csv* yang dapat ditunjukkan pada Gambar 4.



Gambar 4 Hasil *Scraping Reviews Google Meet Appfollow*

Selanjutnya pada Gambar 5 ditunjukkan hasil dari proses *labelling* dalam format *csv* yang dilakukan secara manual menurut pendapat pribadi penulis.



Gambar 5 Hasil *labelling* Ulasan

2.2 Preprocessing Data

Preprocessing adalah teknik maupun strategi yang bertujuan untuk membuat suatu data lebih mudah untuk dikelola atau cocok untuk digunakan pada *text mining* yang tentunya bertujuan agar meningkatkan hasil dari analisis *text mining* [7]. Adapun beberapa tahapan yang perlu dilakukan yaitu. (1) *Case folding* merupakan proses untuk mengubah huruf besar pada setiap kalimat ulasan menjadi

lowercase serta membersihkan data ulasan dari hal yang tidak diperlukan dalam proses pengolahan seperti tanda baca, spasi berlebih, dan sebagainya. (2) *Tokenizing* merupakan proses yang dilakukan untuk memecah kalimat menjadi kata atau entitas lainnya yang disebut sebagai token. (3) *Filtering* merupakan proses yang dilakukan untuk memilah kata-kata penting, baik membuang kata tidak berpengaruh atau menyimpan kata berpengaruh. (4) *Stemming* merupakan proses yang dilakukan untuk membuat sebuah kata menjadi format kata dasar dengan menghilangkan imbuhan dari depan atau belakang kata tersebut.

2.3 Tf-idf Vectorizer

Tf-idf Vectorizer digunakan dalam mengetahui kemunculan *term* dalam dokumen. Dalam hal *Term Frequency (TF)*, semakin sering sebuah *term* muncul dalam sebuah dokumen, semakin tinggi nilai bobot dari term itu sendiri. Sementara, pada proses *Inverse Document Frequency (IDF)*, semakin tinggi frekuensi kemunculan suatu term, maka semakin rendah nilai bobot dari term itu sendiri [8]. Persamaan (1) menunjukkan formula *tf-idf term weighting* pada *scikit learn* [9].

$$idf(t) = \log \frac{1+n}{1+df(t)} + 1 \quad (1)$$

Keterangan:

n = Jumlah total dokumen dalam kumpulan dokumen.

$df(t)$ = Jumlah dokumen dalam kumpulan dokumen yang mengandung term t .

2.4 Klasifikasi dengan *Naïve Bayes Classifiers*

Naïve Bayes Classifier merupakan salah satu pengklasifikasian *probabilistic* paling sederhana [10]. Algoritma *Naïve Bayes Classifier* dapat menghitung sekumpulan probabilitas dengan menghitung frekuensi dan kombinasi nilai dalam kumpulan data yang diberikan [11]. Teorema *bayes* adalah dasar aturan dari *Naïve Bayes Classifier*, berikut teorema *bayes* yang akan ditunjukkan dalam bentuk persamaan (2) [2].

$$P\{X|H\} = \frac{P(X|H).P(H)}{P(X)} \quad (2)$$

Keterangan:

X = Data *tuple* hasil pengujian dari suatu set data yang telah ditentukan masuk ke dalam kelas tertentu.

H_2 = Hipotesis yang akan menentukan X masuk ke dalam kelas C.

$P(H|X)$ = Peluang atau probabilitas X yang merupakan data *tuple* atau bukti yang diperoleh pada saat observasi masuk ke dalam kelas C, probabilitas posterior dimana H dikondisikan pada X.

$P(H)$ = Probabilitas prior, atau probabilitas sebelumnya.

$P(X|H)$ = Probabilitas posterior dimana X dikondisikan pada H.

$P(X)$ = Probabilitas sebelumnya dari X.

2.5 Validasi

K-Fold Cross-Validation merupakan suatu metode yang digunakan sebagai validasi dengan melakukan pembagian data ke dalam k-subset, yang kemudian dilakukan pengulangan sebanyak k kali untuk pembelajaran dan pengujian. Setiap pengulangan, digunakan satu subset sebagai data uji dan subset lainnya sebagai data pembelajaran [13]. Parameter pengujian yang digunakan sebagai evaluasi dari *K-Fold Cross-Validation* yaitu akurasi. Metode pengukuran yang digunakan adalah *confusion matrix*.

2.6 Teks Association

Teks Association adalah metode data mining yang dapat menemukan interkoneksi asosiasi antar item data dalam suatu transaksi data [14]. Langkah terpenting yang diperlukan untuk aturan asosiasi adalah mengetahui seberapa sering kombinasi item terjadi dalam database. Ini disebut sebagai *frequent patterns*. Penting tidaknya aturan asosiatif dapat diketahui dengan dua parameter, *support* yaitu persentase kombinasi item dalam *database* dan *confidence* yaitu kuatnya hubungan antar item dalam aturan asosiatif [15].

3. HASIL DAN PEMBAHASAN

3.1 Data Scraping

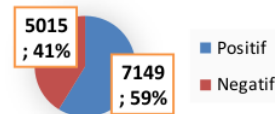
Pada penelitian ini digunakan data hasil *scraping* ulasan *Google Meet* yang ditambang menggunakan teknik *web scraping* pada *web application* yaitu *Appfollow.io*. Data ulasan yang telah berhasil ditambang adalah sebanyak 12.164 data dalam periode waktu 1 Januari 2021 sampai dengan 25 Agustus 2021, yang berhasil di *export* dalam format file *csv*.

3.2 Pelabelan Sentimen

Setelah data berhasil dikumpulkan selanjutnya dilakukan proses pelabelan kelas sentimen berdasarkan respon pengguna terhadap aplikasi *Google Meet* melalui ulasan pada situs *Google Play*. Pelabelan dibagi menjadi dua kelas yaitu kelas positif dan negatif yang dilakukan secara manual menurut pandangan pribadi penulis. Rincian hasil pelabelan sentimen dapat ditunjukkan pada Tabel 1.

Tabel 1 Hasil Pelabelan Sentimen

Sentimen	Frekuensi	Persentase
Positif	7.149	58,8 %
Negatif	5.015	41,2 %



Gambar 6 Diagram Hasil Pelabelan Sentimen

Berdasarkan diagram pada Gambar 6 dapat disimpulkan bahwa ulasan pengguna *Google Meet* dalam periode 1 Januari 2021 sampai dengan 25 Agustus 2021 lebih banyak mengandung sentimen positif. Frekuensi sentimen positif, sejumlah 7.149 ulasan, lebih tinggi dari frekuensi sentimen negatif, sejumlah 5.015 ulasan. Hingga 59% ulasan berisi sentimen positif terkait kinerja aplikasi *Google Meet*. Ini berarti bahwa sebagian besar pengguna cenderung bereaksi positif, memiliki respon yang baik terhadap kinerja aplikasi atau memberikan apresiasi terhadap aplikasi. 41% ulasan berisi sentimen negatif terkait aplikasi *Google Meet*. Dengan kata lain, pengguna menunjukkan respon berupa kritik atau saran terhadap kinerja aplikasi *Google Meet*.

3.3 Preprocessing Data

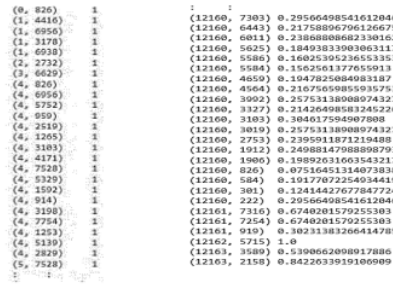
Data yang akan digunakan dalam proses *preprocessing* merupakan data hasil proses *labelling*. Dengan menggunakan *tools Google Colab* tahapan *preprocessing* yang akan dilakukan yaitu *case folding*, *tokenizing*, *stopword removal*, dan *stemming*. Ditunjukkan proses dari tahapan *preprocessing* data pada Tabel 2 berikut.

Tabel 2 Proses Tahapan Preprocessing Data

Tahapan	Input	Output
Case Foldin g	aplikasi yang sangat membantu sekali buat tatap muka sambil belajar 😊😊😊😊😊	aplikasi yang sangat membantu sekali buat tatap muka sambil belajar δŸŸ°δŸŸ°δŸŸ°δŸŸ° ŸŸ°δŸŸ°
Token izing	aplikasi yang sangat membantu sekali buat tatap muka sambil belajar δŸŸ°δŸŸ°δŸŸ°δŸŸ° ŸŸ°δŸŸ°	['aplikasi', 'yang', 'sangat', 'membantu', 'sekali', 'buat', 'tatap', 'muka', 'sambil', 'belajar', 'δŸŸ°δŸŸ°δŸŸ°δŸŸ° ŸŸ°δŸŸ°']
Stopw ord remov al	['aplikasi', 'yang', 'sangat', 'membantu', 'sekali', 'buat', 'tatap', 'muka', 'sambil', 'belajar', 'δŸŸ°δŸŸ°δŸŸ°δŸŸ° δŸŸ°δŸŸ°']	['aplikasi', 'membantu', 'tatap', 'muka', 'belajar', 'δŸŸ°δŸŸ°δŸŸ°δŸŸ° δŸŸ°δŸŸ°']
Stemm ing	['aplikasi', 'membantu', 'tatap', 'muka', 'belajar', 'δŸŸ°δŸŸ°δŸŸ°δŸŸ° δŸŸ°δŸŸ°']	aplikasi bantu tatap muka belajar

3. 4 Tf-idf Vectorizer

Setelah melakukan tahapan preprocessing data, hasil dari preprocessing selanjutnya akan dilakukan perhitungan untuk pembobotan term. Pembobotan term dilakukan dengan menggunakan TfidfVectorizer dari library sklearn. Prosesnya akan dilakukan import fungsi TfidfVectorizer dari sklearn.feature_extraction.text, lalu akan dilakukan inialisasi vectorizer kemudian memanggil fit lalu mengubahnya untuk menghitung skor tf-idf untuk teks. Dan selanjutnya, sklearn fit_transform akan menjalankan fungsi fit dan transform. Hasil akhir yang diperoleh dalam proses tf-idf ditampilkan dalam bentuk skewed matrix yang ditunjukan pada Gambar 7.



Gambar 7 Skewed Matrix Tf-idf Vectorizer

Dari hasil proses tf-idf vectorizer pada Gambar 7 dapat dijelaskan bahwa, pada kolom pertama (0, 826), bilangan (0) merepresentasikan index dari dokumen pada corpus, artinya index ke-0 merepresentasikan kalimat pertama dalam corpus, sedangkan bilangan (826) merepresentasikan index dari token yang terdapat dalam kalimat tersebut, artinya kalimat pada index ke-0 mengandung kata dalam token pada index ke-826. Selanjutnya arti dari sebuah atau sekumpulan angka pada kolom ke-2 merepresentasikan bobot dari tf-idf hasil dari kalkulasi yang dilakukan oleh TfidfVectorizer.

3. 5 Pemodelan Klasifikasi dengan Naïve Bayes Classifiers

Dalam proses pemodelan klasifikasi menggunakan algoritma Naïve Bayes Classifiers, sebelumnya akan dilakukan proses pembagian data latih dan data uji. Probabilitas data dihitung berdasarkan pelatihan yang dilakukan, dan model dibangun dari data uji. Pada penelitian ini dilakukan perbandingan data latih sebesar 20%, data uji sebesar 80%. Yaitu, 9.731 data latih dan 2.432 data uji. Kemudian dilakukan pengukuran terhadap ketepatan dalam klasifikasi dengan confusion matrix berdasarkan hasil dari proses prediksi yang ditunjukkan dalam Tabel 3.

Tabel 3 Hasil Confusion Matrix

Sentimen	Frekuensi	
	Negatif	Positif
Negatif	788	232
Positif	117	1296

Tabel 3 menunjukkan hasil prediksi dari 1.020 data ulasan yang diklasifikasikan dalam kategori sentimen. 788 data ulasan diklasifikasikan dengan benar sebagai true

negative, dan 232 data ulasan salah diklasifikasikan sebagai *false positive*. Sedangkan dari 1.413 data ulasan yang masuk dalam kategori sentimen positif dan negatif, sebanyak 1.296 data komentar diklasifikasikan dengan benar sebagai *true positive* dan 1.117 data komentar tidak diklasifikasikan dengan benar yaitu sebagai *false negative*. Selanjutnya dilakukan perhitungan ketepatan *accuracy*, *precision*, *recall*, dan *f1_score* dalam klasifikasi yang dapat ditunjukkan pada Tabel 4.

Tabel 4 Hasil Perhitungan Ketepatan

Performa	Nilai
<i>Accuracy</i>	0,8565
<i>Precision</i>	0,8701
<i>Recall</i>	0,7725
<i>F1_score</i>	0,8187

Berdasarkan Tabel 4 pemodelan dengan *Naïve Bayes Classifiers* mendapatkan akurasi sebesar 85,65 %. Oleh karena itu, model ini cukup baik dalam mengklasifikasikan ulasan pengguna *Google Meet* pada situs *Google Play*.

3. 6 Validasi dengan *K-Fold Cross-Validation*

Dalam memastikan hasil uji dilakukan pengujian ulang menggunakan *K-Fold Cross-Validation* untuk menilai kinerja proses pada metode algoritma *Naïve Bayes Classifiers*. Dalam penelitian ini jumlah nilai *K* yang digunakan adalah 10 sehingga memiliki 10 *subset* data. Seluruh *dataset* diacak dan dibagi menjadi *K* bagian yang sama. Di sini, kombinasi pertama dari sembilan himpunan bagian yang berbeda digabungkan dan digunakan sebagai data pelatihan, dan bagian yang tersisa digunakan sebagai data uji. Setelah itu, proses pelatihan dan pengujian berlanjut hingga *fold* ke-10. Hasil dari pengujian dapat ditunjukkan dalam Tabel 5.

Tabel 5 Hasil Pengujian 10-Fold Cross-Validation

Iterasi-ke	Akurasi dalam (%)
1	82%
2	84%
3	87%
4	88%

5	89%
6	87%
7	81%
8	85%
9	86%
10	87%
Rata-rata	85,48%

Berdasarkan Tabel 5, hasil akurasi keseluruhan untuk proses *10-Fold Cross-Validation* menunjukkan bahwa nilai akurasi untuk setiap *fold* sangat stabil, dengan sedikit perbedaan nilai pada 81% hingga 89% untuk setiap *fold*-nya. Dalam hal ini percobaan *fold* ke-7 menunjukkan hasil terendah (81%) dari 9 *fold* lainnya, dan percobaan *fold* ke-5 menunjukkan hasil tertinggi (89%) dari 9 *fold* lainnya. Analisis kinerja klasifikasi yang dibuat menggunakan pengklasifikasi *Naïve Bayes* termasuk baik mengingat rata-rata keseluruhan adalah 85,48%.

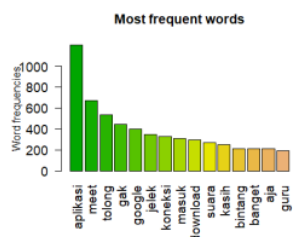
3. 7 Visualisasi

Proses visualisasi data dilakukan untuk mengetahui informasi terkait topik bahasan yang sering diulas oleh pengguna aplikasi *Google Meet* pada situs *Google Play*.



Gambar 8 Wordcloud Ulasan Negatif

Berdasarkan *wordcloud* pada Gambar 8, terlihat kata “aplikasi”, “meet” dan “tolong” adalah kata yang paling dominan dibandingkan dengan kata lain. Banyak orang menggunakan istilah ini dalam isi ulasan mereka mengenai *reviews* dari *Google Meet*. Hasil dari *Wordcloud* merupakan format deskriptif umum, untuk dapat mengetahui informasi yang lebih rinci dapat ditunjukkan dalam Gambar 9 berikut.



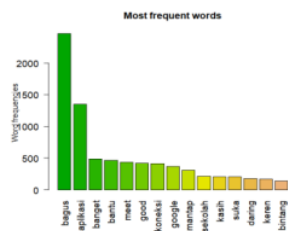
Gambar 9 Diagram *Most Frequent Word Negative*

Gambar 9 menunjukkan *Most Frequent Word* pada ulasan negatif. Hasil yang diperoleh searah dengan *Wordcloud* dimana "aplikasi", "meet" dan "tolong" merupakan kata yang paling dominan. Kata "aplikasi" muncul sebanyak 1199 kali, diikuti oleh "meet" 673 kali, dan "tolong" 536 kali. Juga selain itu frekuensi kata yang sering muncul adalah "jelek", "koneksi", "masuk", "download", dan "suara". Hal tersebut menunjukkan bahwa ulasan negatif pada *Google Meet* membahas terutama mengenai masalah performa dari fitur pada aplikasi. Dalam hal ini beberapa dari pengguna menganggap performa aplikasi kurang baik karena mengandung masalah seperti koneksi yang menurun, seringkali mengalami kesulitan dan gagal dalam melakukan *download* atau masuk pada aplikasi, serta suara yang seringkali tersendat, menggeema, dan memantul.



Gambar 10 *Wordcloud Ulasan Positif*

Berbeda dengan *wordcloud* ulasan negatif pada Gambar 10, terlihat kata "bagus", "aplikasi" dan "banget" adalah kata yang paling dominan dibandingkan dengan kata lain. Hasil dari *Wordcloud* tersebut merupakan format deskriptif umum. Untuk mengetahui informasi yang lebih rinci dapat ditunjukkan dalam Gambar 11 berikut.



Gambar 11 *Most Frequent Word Positive*

Gambar 11 menunjukkan *Most Frequent Word* pada ulasan positif. Hasil yang diperoleh seiring dengan *Wordcloud* dimana "bagus", "aplikasi" dan "banget" merupakan kata yang paling dominan. Kata "bagus" muncul sebanyak 2404 kali, diikuti oleh "aplikasi" 1361 kali, dan "banget" 491 kali. Selanjutnya selain itu frekuensi kata yang sering muncul adalah "bantu", "good", "koneksi", "mantap", "sekolah", dan "daring". Hal tersebut menunjukkan bahwa ulasan positif pada *Google Meet* membahas terutama mengenai masalah performa fitur dari aplikasi yang sudah cukup baik. Dalam hal ini dapat dikatakan aplikasi memiliki kinerja baik yang dapat dipertahankan dan ditingkatkan serta sangat membantu bagi pengguna dalam melakukan kegiatan daring.

3.8 Text Association

Text Association dilakukan untuk menemukan asosiasi yang paling sering terjadi antar kata pada saat yang sama. Hal ini dapat memperkuat dalam melakukan prediksi ketika sedang melakukan pencarian informasi pada suatu sebab topik masalah yang sedang dibicarakan baik pada ulasan negatif maupun positif. Dalam hal ini *teks Association* dilakukan dengan bantuan *tools R*. Hasil dari asosiasi dapat ditunjukkan dalam Tabel 6 dan 7.

Tabel 6 *Text Association Ulasan Negatif*

Koneksi		Audio	
Lambat	0,20	Musik	0,26
Telkomsel	0,20	Mute	0,21
Patah	0,20	Putus	0,19
		Disconnect	0,19

Video		Masalah	
Quality	0,27	Downloaded	0,30
Blur	0,26	Sendat	0,26
Resolution	0,17	Rapat	0,23
Jeda	0,17	Mikrofon	0,17
Delay	0,17	Handphone	0,17
Buram	0,17		
Interface	0,16		
Suara		Camera	
Putus	0,29	Payah	0,24
Gema	0,24	Mirror	0,24
Pantul	0,21	Setting	0,22
Kualitas		Fitur	
Design	0,33	Tambah	0,26
Forced	0,20	Background	0,19
Buruk	0,16	Filter	0,16
Lamban	0,16	Fullscreen	0,16
Gambar	0,16	Resolusi	0,15
Keluar	0,16		
Update			
Kendala	0,22		
Ngebug	0,20		

Berdasarkan Tabel 6 diperoleh beberapa asosiasi kata yang berhubungan dengan kata “koneksi” memiliki asosiasi dengan kata “lambat” dan “patah” sehingga pengguna mengalami kendala sendatan pada saat melakukan *meeting*. Untuk kata “audio” memiliki asosiasi dengan kata “mute”, “putus” dan “disconnect” dimana seringkali audio mengalami kendala putus-putus, *disconnect*, bahkan tidak terdengar oleh peserta rapat. kata “video” memiliki asosiasi dengan kata “quality”, “blur”, “resolusi”, “interface”, dan “jeda” dimana pengguna menimbang kualitas dari video yang ditampilkan sering mengalami blur atau buram, memiliki resolusi dan *interface* yang kurang cakap serta sering mengalami jeda. kata “masalah” memiliki asosiasi dengan kata “download”, dan “microfon” dimana saat melakukan *download* sangat diperlukan waktu yang lama lalu tiba-tiba gagal terunduh. Beberapa juga mengalami kendala suara dari mikrofon yang sering kali tidak terdengar terhadap sesama peserta *meeting* sementara mikrofon sudah dinyalakan. kata “suara” memiliki asosiasi dengan kata “putus”, “gema”, “pantul” sehingga pada saat tertentu suara sering mengalami putus, memiliki gema dan pantulan jika ada *device* yang saling berdekatan di lingkungan yang sama sehingga hal ini kerap membuat pengguna kurang nyaman saat menggunakan *Google Meet*. kata “camera” memiliki asosiasi dengan

kata “payah”, “mirror”, dan “setting” dimana pengguna menilai *camera* cukup gelap, mengalami error seperti hitam pada layar bahkan ketika *camera* telah dinyalakan, serta ketiadaan fitur *mirror* dan *setting* sehingga diperlukan pengembangan. kata “kualitas” memiliki asosiasi dengan kata “lamban”, “design”, dan “keluar” dimana aplikasi berjalan dengan lamban, *design interface* baru yang dirasa pengguna lebih rumit dibandingkan dengan versi terdahulu, serta aplikasi yang sering *forced close* atau keluar dengan sendirinya. kata yang berhubungan dengan kata “fitur” memiliki asosiasi dengan kata “background”, “filter”, “fullscreen” dan “resolusi” sehingga pengguna mengharapkan penambahan fitur guna meningkatkan kenyamanan dalam penggunaan aplikasi. kata yang berhubungan dengan kata “update” memiliki asosiasi dengan kata “kendala” dan “ngebug” sehingga pengguna menimbang *update* yang dilakukan oleh *Google Meet* justru banyak menambah kendala pada aplikasi. Diharapkan pada *update* selanjutnya pihak aplikasi dapat memperbaiki *bug* dan kendala yang ditimbulkan.

Tabel 7 Text Association Ulasan Positif

Meeting		Kualitas	
Kantor	(0,29)	Audio	(0,32)
School	(0,25)	Kencang	(0,32)
Education	(0,23)	Bersih	(0,23)
Akun	(0,18)	Gambar	(0,16)
Kerja	(0,16)	Recommended	(0,16)
Join	(0,16)	Sendat	(0,16)
		Handal	(0,16)
Koneksi		Suara	
Proses	(0,25)	Dengar	(0,28)
Daring	(0,22)	Putus	(0,27)
Mudah	(0,16)	Jernih	(0,26)
Audio		Video	
Cohost	(0,30)	Sidang	(0,30)
Performa	(0,30)	Binar	(0,22)
Pertahankan	(0,30)	Wifi	(0,22)
Microfon	(0,21)	Ngelag	(0,22)
Off	(0,21)	Conference	(0,20)
Quality	(0,21)		

Kamera		Fitur	
Cerah	(0,43)	Camera	(0,41)
Mirror	(0,33)	Mic	(0,47)
Mati	(0,24)	Resolusi	(0,28)
Hitam	(0,24)	Android	(0,24)
Opsi	(0,17)		
Problem			
Fitur	(0,58)		
Perfect	(0,24)		

Berdasarkan Tabel 7 diperoleh beberapa asosiasi kata yang berhubungan dengan kata “meeting” memiliki asosiasi dengan kata “kantor”, “school”, sehingga aplikasi bermanfaat dalam kegiatan produktivitas sehari-hari baik *meeting* dalam urusan kantor maupun sekolah dan dalam bidang pekerjaan maupun edukasi. Untuk kata “kualitas” memiliki asosiasi dengan kata “recommended” dan “handal” dimana pengguna merasa puas dengan performa aplikasi sehingga pengguna merekomendasikan aplikasi *google meet* dalam melakukan aktivitas *meeting*. kata “koneksi” memiliki asosiasi dengan kata “mudah” dan “daring” sehingga pengguna mudah untuk terhubung dalam kegiatan daring, dan pengguna tidak menjumpai masalah yang berarti dalam koneksi. kata “suara” memiliki asosiasi dengan kata “dengar”, dan “jernih” dimana kualitas suara yang dimiliki oleh *Google Meet* sudah cukup jernih. kata “audio” memiliki asosiasi dengan kata “performa”, dan “co-host” sehingga pengguna merasa kualitas serta performa dari *audio Google Meet* layak dipertahankan. Penambahan fitur *co-host* untuk mengatur audio seluruh partisipan *meeting* juga diharapkan. kata “video” memiliki asosiasi dengan kata “sidang”, “binar”, dan “conference” sehingga aplikasi *Google Meet* merupakan aplikasi *video conference* dengan performa *video* yang baik dalam pengalaman beberapa pengguna ketika melakukan kegiatan webinar dan sidang. kata “kamera” memiliki asosiasi dengan kata “cerah”, “mirror”, dan “opsi” dimana pengguna merasa *output* kamera sudah cukup baik. Pengguna berharap pihak aplikasi dapat meningkatkan performa dengan menambah fitur opsi dan *mirror* pada kamera. kata yang berhubungan dengan kata “fitur” memiliki asosiasi dengan kata “camera”, “mic”, dan “resolusi” sehingga pengguna mengharapkan peningkatan performa pada fitur *camera*, *microphone*, dan resolusi. kata yang

berhubungan dengan kata “problem” memiliki asosiasi dengan kata “fitur” dan “perfect” sehingga pengguna tidak menemukan adanya masalah yang terlihat serius. Menurut pengguna aplikasi sudah baik dan memiliki fitur pendukung untuk melakukan *meeting*.

4. KESIMPULAN

Dari penelitian yang telah dilakukan, dapat disimpulkan bahwa berdasarkan 12.164 data. Pengujian klasifikasi sentimen kualitas layanan menggunakan metode *Naïve Bayes Classifier* pada aplikasi *Google Meet* menunjukkan 58,8% sentimen positif dan 41,2% sentimen negatif, sehingga secara inklusif aplikasi *Google Meet* berfungsi dengan baik dan efektif untuk kegiatan daring. Algoritma *Naïve Bayes Classifiers* memiliki kinerja yang baik dalam klasifikasi sentimen, hal ini dibuktikan dengan nilai akurasi sebesar 85,68% dan akurasi rata-rata keseluruhan 85,48% untuk pengujian dengan *10-Fold Cross-Validation*. Berdasarkan penerapan metode *Association*, hasil yang diperoleh untuk ulasan kelas positif adalah ekstraksi informasi terkait *meeting*, *video*, *problem*, koneksi, suara, *audio*, kualitas, kamera, dan fitur. Sedangkan pada ulasan kelas negatif, metode *Association* menunjukkan ekstraksi informasi terkait hal-hal yang sering dikeluhkan di kelas negatif yaitu *video*, masalah, koneksi, *audio*, suara, *update*, kualitas, *camera*, dan fitur.

5. SARAN

Dari penelitian yang telah dilakukan, saran yang dapat diberikan untuk pengembangan dan perbaikan penelitian selanjutnya yaitu meningkatkan kualitas data *preprocessing* dengan teknik *slang word replacing*, *misspelling* dan memperluas list *stopword*, serta disarankan menggunakan metode klasifikasi lain seperti *support vector machine*, *logistic regression*, atau *lexicon-based* sebagai perbandingan untuk memahami pendekatan mana yang lebih baik.

DAFTAR PUSTAKA

- [1] M. A. Pratama, “Analisis Perbaikan Kualitas Layanan Pada Aplikasi Zoom Cloud Meetings Berdasarkan Ulasan

- End-User Menggunakan Metode Text Mining Dan Fishbone Diagram,” Universitas Islam Indonesia, 2021.
- [2] N. Helmiah *et al.*, “Penerapan Metode Naïve Bayes dalam Analisis Persepsi Masyarakat mengenai Rencana Pengesahan RUU Omnibus Law di Bidang Investasi dan Ketenagakerjaan Tahun 2020 di Indonesia,” *J. MSA (Mat. dan Stat. serta Apl.)*, vol. 8, no. 2, p. 48, 2020, doi: 10.24252/msa.v8i2.16743.
- [3] E. B. Santoso and A. Nugroho, “Analisis Sentimen Calon Presiden Indonesia 2019 Berdasarkan Komentar Publik Di Facebook,” *Eksplora Inform.*, vol. 9, no. 1, pp. 60–69, 2019, doi: 10.30864/eksplora.v9i1.254.
- [4] R. Wati, “Penerapan Algoritma Genetika Untuk Seleksi Fitur Pada Analisis Sentimen Review Jasa Maskapai Penerbangan,” *J. Evolusi*, vol. 4, no. 1, pp. 25–31, 2016.
- [5] R. W. Pratiwi, D. Yusuf, and S. Nugroho, “Prediksi Rating Film Menggunakan Metode Naive Bayes,” *J. Tek. Elektro*, vol. 8, no. 2, pp. 60–63, 2016, [Online]. Available: <https://journal.unnes.ac.id/nju/index.php/jte/article/view/7764>.
- [6] V. Oktaviani, B. Warsito, H. Yasin, R. Santoso, and Suparti, “Sentiment analysis of e-commerce application in Traveloka data review on Google Play site using Naïve Bayes classifier and association method,” *J. Phys. Conf. Ser.*, vol. 1943, no. 1, 2021, doi: 10.1088/1742-6596/1943/1/012147.
- [7] A. M. F. Hulu and K. M. Lhaksana, “Analisis Sentimen Politik pada Twitter Menggunakan Metode Support Vector Machine (Studi Kasus : Pilpres 2019),” *Semin. Nas. Teknol. Fak. Tek. Univ. Krisnadwipayana*, vol. 1, no. 1, pp. 739–742, 2019, [Online]. Available: <https://jurnal.teknikunkris.ac.id/index.php/semnastek2019/article/view/343/342>.
- [8] H. Zakiyudin and A. Apriani, “Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF System Penerimaan Mahasiswa Baru pada Kampus Swasta,” *J. Bumigora Inf. Technol.*, vol. 3, no. 1, pp. 19–27, 2021, doi: 10.30812/bite.v3i1.1110.
- [9] S. Developers, “6.2. Feature extraction — scikit-learn 1.0.2 documentation.” https://scikit-learn.org/stable/modules/feature_extraction.html#text-feature-extraction (accessed Apr. 17, 2022).
- [10] S. Taheri and M. Mammadov, “Learning the naive bayes classifier with optimization models,” *Int. J. Appl. Math. Comput. Sci.*, vol. 23, no. 4, pp. 787–795, 2013, doi: 10.2478/amcs-2013-0059.
- [11] M. M. Saritas and A. Yasar, “Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification,” *Int. J. Intell. Syst. Appl. Eng.*, pp. 88–91, 2019, [Online]. Available: <http://xlink.rsc.org/?DOI=C5TC02043C>.
- [12] F. Handayani and S. Pribadi, “Implementasi Algoritma Naive Bayes Classifier dalam Pengklasifikasian Teks Otomatis Pengaduan dan Pelaporan Masyarakat melalui Layanan Call Center 110,” *J. Tek. Elektro*, vol. 7, no. 1, pp. 19–24, 2015, doi: 10.15294/jte.v7i1.8585.
- [13] A. Widjaya, L. Hiryanto, and T. Handhayani, “Prediksi Masa Studi Mahasiswa Dengan Voting Feature Interval 5 Pada Aplikasi Konsultasi Akademik Online,” *J. Comput. Sci. Inf. Syst.*, vol. 1, no. 1, p. 25, 2017, doi: 10.24912/computatio.v1i1.238.
- [14] M. Brilliant, D. Handoko, and Sriyanto, “Implementation of Data Mining Using Association Rules for Transactional Data Analysis,” *3rd Int. Conf. Inf. Technol. Bus.*, pp. 177–180, 2017.
- [15] W. Aprianti, K. A. Hafizd, and M. R. Rizani, “Implementasi Association Rules dengan Algoritma Apriori pada Dataset Kemiskinan,” *Limits J. Math. Its Appl.*, vol. 14, no. 2, p. 57, 2017, doi: 10.12962/limits.v14i2.2933.

Naskah Publikasi-Fina Khoirunnisa-18.01.55.0044-01082022

ORIGINALITY REPORT

3%

SIMILARITY INDEX

2%

INTERNET SOURCES

3%

PUBLICATIONS

3%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to Universitas Sembilanbelas
November Kolaka

Student Paper

2%

2

ejournal.unsri.ac.id

Internet Source

2%

Exclude quotes On

Exclude matches < 2%

Exclude bibliography On