

Analisis Review Pengguna Google Meet dan Zoom Cloud Meeting Menggunakan Algoritma Naïve Bayes

Muhammad Rezki¹, Desiana Nur Kholifah², Muhammad Faisal³, Priyono⁴, Rachmat Suryadithia⁵

^{1,2,3,4,5}Universitas Bina Sarana Informatika

e-mail: ¹muhammad.mdk@bsi.ac.id, ²desiana.dfh@bsi.ac.id, ³muhammad.mal@bsi.ac.id,

⁴Priyono.pyo@bsi.ac.id, ⁵rachmat.rcs@bsi.ac.id

Diterima	Direvisi	Disetujui
09-11-2020	15-11-2020	15-12-2020

Abstrak - Saat ini seluruh dunia sedang menghadapi wabah penyakit menular yaitu virus Covid 19. Pembatasan sosial atau menjaga jarak adalah serangkaian tindakan pengendalian infeksi nonfarmasi yang dimaksudkan untuk menghentikan atau memperlambat penyebaran penyakit menular tersebut. Sehingga seluruh masyarakat diharapkan untuk beraktifitas dirumah untuk menghentikan penyebaran virus Covid 19. Agar tetap bisa menjalankan aktifitas dirumah diperlukan *virtual meet* untuk berkomunikasi sesama team atau rekan kerja. Saat ini *virtual meet* telah banyak dipakai. Penilaian Sebuah Aplikasi di *Playstore* memiliki tujuan untuk memberikan ulasan tentang kelebihan dan kekurangan dalam penggunaan aplikasi khususnya *virtual video conference*. Untuk mengetahui sejauh mana analisa review pengguna aplikasi Google Meet dan Zoom Cloud Meeting berdasarkan pemberian jumlah bintang dengan menggunakan teknik klasifikasi yaitu perbandingan Algoritma *Naïve Bayes* dengan *feature* optimasi *SMOTE Upsampling* dan *PSO*. Penggunaan *feature selection synthetic minority over-sampling technique (SMOTE)* dan *feature optimasi Particle swarm optimization (PSO)* pada algoritma klasifikasi terbukti sangat berpengaruh untuk meningkatkan akurasi pada algoritma *Naïve Bayes* untuk pengolahan data review pengguna Google Meet dan Zoom Cloud Meeting pada google play berdasarkan perolehan skor bintang. Hasil pengujian mendapatkan hasil akurasi sebesar 85,76 % yang ditambah dengan *Feature Smote* dan *PSO* untuk review Zoom Cloud Meeting, sedangkan untuk pengguna Google Meet yang ditambah dengan *Feature Smote* dan *PSO* hanya mampu mendapat tingkat akurasi sebesar 79,33 %.

Kata Kunci: Naive Bayes, Covid-19, Playstore

Abstract - Currently the whole world is facing an outbreak of an infectious disease, namely the Covid 19 virus. Social restrictions or distancing are a series of non-pharmaceutical infection control measures intended to stop or slow the spread of the infectious disease. So that all people are expected to do activities at home to stop the spread of the Covid 19 virus. In order to continue to be able to carry out activities at home, a virtual meet is needed to communicate with fellow teams or colleagues. Currently, virtual meet has been widely used. Assessment of an application in Playstore aims to provide an overview of the advantages and disadvantages of using applications, especially virtual video conferencing. To find out the extent to which the analysis of user reviews of the Google Meet and Zoom Cloud Meeting applications is based on the number of stars by using a classification technique, namely the comparison of the Naïve Bayes Algorithm with the SMOTE Upsampling and PSO optimization features. The use of the synthetic minority over-sampling technique (SMOTE) feature selection and the Particle swarm optimization (PSO) optimization feature in the classification algorithm has proven to be very influential in improving the accuracy of the Naïve Bayes algorithm for processing Google Meet and Zoom Cloud Meeting user review data on Google Play based on acquisition. star score. The test results get an accuracy of 85.76% which is added with the Smote and PSO Feature for the Zoom Cloud Meeting review, while for Google Meet users who are added with the Smote and PSO Features are only able to get an accuracy rate of 79.33%.

Keywords: Naive Bayes, Covid-19, Playstore

PENDAHULUAN

Saat ini seluruh adunia sedang menghadapi wabah penyakit menular yaitu virus Covid 19. Pembatasan sosial atau menjaga jarak adalah

serangkaian tindakan pengendalian infeksi nonfarmasi yang dimaksudkan untuk menghentikan atau memperlambat penyebaran penyakit menular tersebut. Sehingga seluruh masyarakat diharapkan untuk beraktifitas dirumah untuk menghentikan

penyebaran virus Covid 19. Agar tetap bisa menjalankan aktifitas dirumah diperlukan *virtual meet* untuk berkomunikasi sesama team atau rekan kerja. Saat ini *virtual meet* telah banyak dipakai.

Google Play adalah layanan konten digital milik Google yang terdiri dari toko produk-produk online seperti musik/lagu, buku, aplikasi, permainan, ataupun pemutar media. Google Play dapat diakses melalui web, aplikasi android (PlayStore), dan Google TV. Dalam Google Play, selain *virtual meet* juga dilengkapi adanya fitur penilaian bagi pelanggan untuk memberikan ulasan tentang kelebihan dan kekurangan dalam penggunaan aplikasi *virtual video conference*, ulasan pengguna bisa bermacam-macam bentuknya, mulai dari kalimat yang halus maupun kasar juga ada tergantung penilaian masing-masing penggunanya. Pengguna google play, selain memberikan ulasan juga bisa memberikan penilaian berupa skor bintang (antara 1 – 5) tergantung penilaian yang diberikan oleh pengguna tersebut dan terakhir pengguna juga bisa memberikan skor “like” terhadap komentar seseorang yang memang bisa mewakili ketertarikan atau ketidaksukaan seseorang terhadap aplikasi *virtual meet* yang digunakan.

Ulasan dari pengguna sering digunakan sebagai alat yang efektif dan efisien dalam menemukan informasi terhadap suatu produk atau jasa. Bahwa kebanyakan dari pengguna internet bergantung pada rekomendasi *word-of-mouth* (opini) sebelum menggunakan suatu produk, karena *review* dari pengguna lain dapat menyediakan informasi terbaru dari produk tersebut berdasarkan perspektif pengguna-pengguna lain yang sudah menggunakan produk tersebut.

Untuk penelitian ini, fokus utamanya adalah *review* pengguna aplikasi Zoom Cloud Meeting dan Google Meet yang ada di situs Google Play dimana untuk pengambilan datanya pada Bulan Maret 2020 – Mei 2020 (kurun waktu 3 bulan). Citra merek yang baik tentunya akan membentuk opini yang baik pula dari konsumen tentang suatu produk/jasa. *Review* pengguna Zoom Cloud Meeting maupun Google Meet bisa jadi dipengaruhi oleh beberapa hal yang belum menjadi perhatian baik dari pihak Zoom Cloud Meeting maupun Google Meet. Hal ini mungkin terjadi karena adanya beberapa faktor yang harus diperbaiki dan belum diketahui oleh pihak Zoom meeting maupun Google Meet, agar kualitas dari aplikasi semakin baik.

Salah satu analisis text mining yaitu analisis sentimen dapat diaplikasikan pada perusahaan yang mengeluarkan produk dan jasa yang menyediakan layanan untuk menerima pendapat (feedback) dari konsumen untuk produk tersebut. Analisis sentimen diaplikasikan untuk mengelompokkan feedback positif, negatif, dan netral dari konsumen sehingga membantu tugas perusahaan untuk meninjau kembali kekurangan produk mereka. Apabila ditemukan adanya sentiment negatif, maka

perusahaan dapat dengan cepat mengambil tindakan untuk memperbaiki produknya. Analisis sentimen merupakan riset komputasional dari opini, sentimen, dan emosi yang diekspresikan secara tekstual (Gunawan et al., 2017).

Terdapat beberapa penelitian yang sudah dilakukan dalam melakukan klasifikasi sentimen terhadap review yang tersedia secara online diantaranya, Analisis Sentimen Pada Ulasan Aplikasi Mobile Menggunakan Naïve Bayes dan Normalisasi Kata Berbasis Levenshtein Distance (Studi Kasus Aplikasi BCA Mobile) (Gunawan et al., 2017). Analisis Sentimen Data Komentar Sosial Media Facebook Dengan K-Nearest Neighbor (Studi Kasus Pada Akun Jasa Ekspedisi Barang J&T Ekspres Indonesia) (Salam et al., 2018). Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexicon Based Features (Rofiqoh et al., 2017). Analisa Sentimen Untuk Penilaian Pelayanan Situs Belanja Online Menggunakan Algoritma Naïve Bayes (Muljono et al., 2018). Analisa Model Support Vector Machine Text mining Pada Komentar Positif Dan Negatif Untuk Review Perbandingan Whatsapp Vs Bbm (Darmawan & Syamsiah, 2016). Aplikasi Web Untuk Analisis Sentimen Pada Opini Produk Dengan Metode Naïve Bayes Classifier (Hanggara et al., 2017). Dari beberapa penelitian terkait, kami mengambil metode Naïve Bayes berbasis feature selection, sehingga penelitian ini bertujuan untuk membandingkan 2 aplikasi virtual meet yaitu Google Meet dan Zoom Cloud Meeting dengan sumber data review pengguna pada google play berdasarkan rating bintang yang didapatkan.

1. Analisis Sentimen

Analisis sentimen atau opinion mining merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini (Buntoro, 2017). Analisis sentimen dapat dibedakan berdasarkan sumber datanya, beberapa level yang sering digunakan dalam penelitian analisis sentimen adalah analisis sentimen pada level dokumen dan analisis sentimen pada level kalimat (Pertiwi, 2019).

Berdasarkan level sumber datanya Analisis Sentimen terbagi menjadi dua kelompok besar yaitu (Hulu & Lhaksana, 2019):

a. Coarse-grained Sentiment Analysis

Pada Sentiment Analysis Coarse-grained, Analisis Sentimen yang dilakukan adalah pada level dokumen. Secara garis besar fokus utama dari Sentiment Analysis jenis ini adalah menganggap seluruh isi dokumen sebagai sebuah sentiment positif atau sentiment negatif.

b. Fined-grained Sentiment Analysis

Fined-grained Sentiment Analysis adalah

Sentiment Analysis pada level kalimat. Fokus utama *fined-grained Sentiment Analysis* adalah menentukan sentimen pada setiap kalimat.

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data besar (Gunadi & Sensuse, 2012). Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi. Data mining adalah kegiatan menemukan pola yang menarik dari data dalam jumlah besar, data dapat disimpan dalam database, data warehouse, atau penyimpanan informasi lainnya. Data mining berkaitan dengan bidang ilmu-ilmu lain, seperti database sistem, data *warehousing*, statistik, *machine learning*, *information retrieval*, dan komputasi tingkat tinggi (Abdillah et al., 2016).

Karakteristik data mining adalah sebagai berikut (Abdillah et al., 2016):

1. Data mining berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.
2. Data mining biasa menggunakan data yang sangat besar. Biasanya data yang besar digunakan untuk membuat hasil lebih dipercaya.
3. Data mining berguna untuk membuat keputusan yang kritis, terutama dalam strategi.

Menurut Larose dalam bukunya yang berjudul “*Discovering Knowledge in Data: An Introduction to Data Mining*”, data mining dibagi menjadi beberapa kelompok berdasarkan tugas/pekerjaan yang dapat dilakukan, yaitu (Gunadi & Sensuse, 2012):

a. Deskripsi

Terkadang peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola kecenderungan sering memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

b. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan baris data (*record*) lengkap yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.

c. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang. Beberapa metode dan teknik yang digunakan dalam klasifikasi dan estimasi dapat pula digunakan (untuk keadaan yang

tepat) untuk prediksi.

d. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

e. Pengklasteran (*Clustering*)

Pengklasteran merupakan pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas obyek-obyek yang memiliki kemiripan. Klaster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidakmiripan *record* dalam klaster yang lain. Berbeda dengan klasifikasi, pada pengklasteran tidak ada variabel target. Pengklasteran tidak melakukan klasifikasi, mengestimasi, atau memprediksi nilai dari variabel target, akan tetapi, algoritma pengklasteran mencoba untuk melakukan pembagian terhadap keseluruhan data menjadi kelompok-kelompok yang memiliki kemiripan (homogen), yang mana kemiripan *record* dalam satu kelompok akan bernilai maksimal, sedangkan kemiripan dengan *record* dalam kelompok lain akan bernilai minimal.

f. Asosiasi

Tugas asosiasi dalam data mining adalah untuk menemukan atribut yang muncul dalam satu waktu. Salah satu implementasi dari asosiasi adalah *market basket analysis* atau analisis keranjang belanja, sebagaimana yang akan dibahas dalam tesis ini.

2. Text Mining

Text mining (menambang teks) merupakan analisis teks dimana sumber data biasanya di dapatkan dari dokumen, dan tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisa keterhubungan, keterkaitan dan kelas antar dokumen (Hartanto, 2017). *Text mining* merupakan salah satu teknik yang dapat digunakan untuk melakukan klasifikasi, dimana *text mining* merupakan variasi dari data mining yang berusaha menemukan pola yang menarik dari sekumpulan data tekstual yang berjumlah besar (Ratniasih et al., 2017). Dalam penerapan *text mining*, terdapat beberapa langkah yang perlu dilakukan antara lain (Ratniasih et al., 2017):

a. Tokenizing

Tokenizing merupakan proses penguraian Deskripsi yang semula berupa kalimat menjadi kata.

b. Filtering

Filtering adalah tahap mengambil kata penting dari hasil proses token. Bisa menggunakan algoritma *stop list* atau *word list*. *Filtering* dapat juga diartikan sebagai proses mengambil kata – kata penting dari hasil proses token atau penghapusan *stopwords*. *Stopwords* merupakan kosa kata yang bukan merupakan ciri (kata unik) dari suatu

dokumen.

c. Stemming

Stemming merupakan tahap untuk mencari root kata dari hasil filtering. Stemming adalah proses pemetaan dan penguraian berbagai bentuk (variants) dari suatu kata menjadi bentuk kata dasarnya (stem).

d. Tagging

Tagging merupakan tahap untuk mencari bentuk awal/root dari tiap kata lampau atau hasil dari proses stemming.

e. Tahap Analysing

Analyzing merupakan tahap penentuan seberapa jauh keterhubungan antar suatu kata atau term terhadap suatu dokumen atau kalimat dengan menghitung nilai/bobot keterhubungan. Algoritma TF/IDF digunakan dalam proses perhitungan bobot (W) terminologi kata. Algoritma ini digunakan untuk menghitung bobot setiap kata yang paling umum digunakan pada information retrieval.

3. Naïve Bayes

Algoritma Naïve Bayes adalah pengelompokan probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menghitung frekuensi dan kombinasi nilai-nilai dalam suatu kumpulan data yang diberikan (Saritas & Yasar, 2019).

Perhitungan Naïve Bayes dapat dilakukan dengan langkah berikut ini (Minarni & Irawan, 2019):

- Mencari nilai prioritas untuk tiap-tiap kelas dengan menghitung rata-rata tiap kelas dengan menggunakan persamaan.
- Mencari nilai Likelihood untuk tiap-tiap kelas dengan menggunakan persamaan.
- Mencari nilai posterior dari tiap kelas yang ada dengan menggunakan persamaan.

METODE PENELITIAN

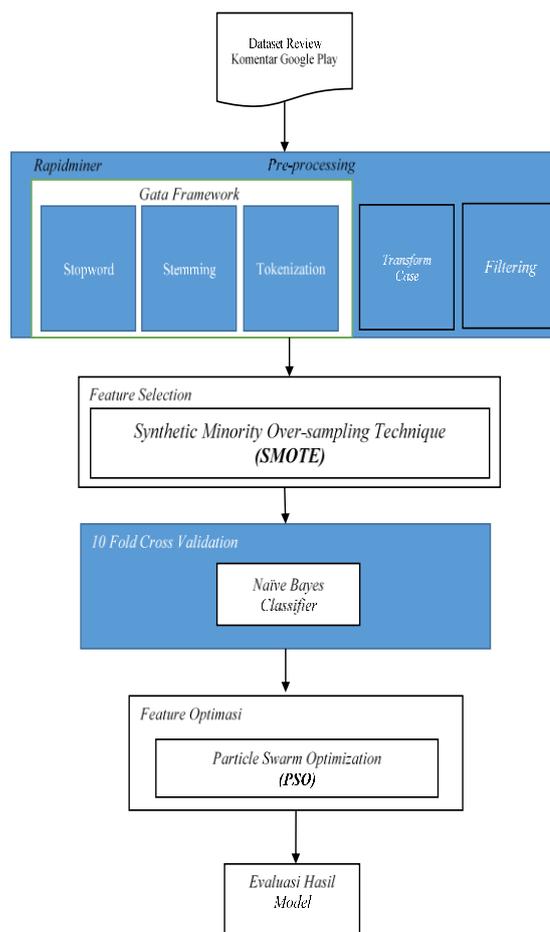
Metodologi sendiri merujuk kepada alur pemikiran umum atau menyeluruh (general logic) dan gagasan teoritis (theoretic perspectives) suatu penelitian. Sedangkan kata 'metode' menunjuk pada teknik yang digunakan dalam penelitian seperti survey, wawancara dan observasi. Suatu penelitian yang baik senantiasa memperhatikan kesesuaian antara teknik yang digunakan dengan alur pemikiran umum serta gagasan teoritis (Raco, 2010).

Penelitian ini menggunakan data kuantitatif untuk menganalisis data dari teks berbahasa Indonesia berupa ulasan pengguna aplikasi Google Meet dan Zoom Cloud meeting yang didapat dari google play, data ulasan pengguna aplikasi Virtual meet baik Google Meet dan Zoom Cloud meeting yang sudah mendapatkan skor penilaian berupa banyaknya bintang yang diberikan oleh pengguna aplikasi tersebut. Dalam tahapan mencari pemecahan masalah ini akan dilakukan dengan kerangka pemikiran menggunakan model penelitian Knowledge Discovery in Database (KDD), dimana

tahapannya meliputi:

- Selection
- Pre-Processing
- Transformation
- Data Mining
- Interpretation/Evaluation

Masing-masing tahapan diatas, akan dijelaskan melalui gambar dibawah ini untuk menjelaskan alur penelitian Analisis Review Pengguna Google Meet dan Zoom Cloud meeting Menggunakan Algoritma Naïve Bayes.



Gambar 1. Kerangka Pemikiran

Sumber: Hasil penelitian (2020)

HASIL DAN PEMBAHASAN

Pada penelitian ini, review pengguna Google Meet dan Zoom Cloud Meeting data diambil dari google play, akan dilakukan pengujian untuk membandingkan perolehan skor penilaian bintang 1 dan bintang 5 dimana jumlahnya sekitar 714 data. Dari review pengguna tersebut, kemudian dikelompokkan berdasarkan skor penilaian bintang yang diberikan oleh pengguna Google Meet dan Zoom Cloud Meeting, dimana untuk gambarannya seperti tabel berikut ini:

Tabel 1. Jumlah Data Review Pengguna Bintang 1 dan 5

No	Review Pengguna	Jumlah Bintang		
		1	5	Total
1	Google Meet	152	254	406
2	Zoom Cloud Meeting	59	249	308
Jumlah				714

Sumber: Telah diolah kembali berdasarkan pencarian di *google play* (2020)

Dalam mengolah data untuk mendapatkan model yang sesuai dengan kasus penelitian ini yaitu Review Analisis Pengguna Google Meet dan Zoom Cloud Meeting menggunakan Algoritma Naïve Bayes digunakanlah sebuah tools Rapidminer versi 9.7. Karena penelitian ini termasuk bagian dari text-mining maka ada tahap yang harus dilakukan terlebih dulu sebelum bisa didapatkan sebuah model yang baik pada studi kasus analisis komentar pengguna Google Meet dan Zoom Cloud Meeting pada google play.

Pembahasan pada tahapan ini adalah proses awal pengolahan dataset sebelum dapat diproses untuk klasifikasi dengan algoritma Naïve Bayes yang dioptimalkan dengan menggunakan feature Synthetic Minority Over-Sampling Technique (SMOTE) dan Particle Swarm Optimization (PSO). Penelitian ini menggunakan beberapa tahapan pre-processing untuk dataset teks komentar, berikut ini adalah tahapan tersebut:

1. Case Folding
2. Tokenizing
3. Filtering
4. Stemming

Tahapan-tahapan pre-processing teks komentar diatas diproses dengan menggunakan tools Gata Framework Textmining yang dapat diakses melalui www.gataframework.com/textmining, untuk menggunakan tools tersebut dapat dalam single teks yang dapat langsung di masukan pada kolom yang tersedia di halaman tools tersebut, namun juga dapat dalam bentuk file excel yang kemudian diunggah.

Gata Framework merupakan alternatif dalam pre-processing teks berbahasa Indonesia yang dikombinasikan dengan aplikasi RapidMiner untuk memproses kata-kata dalam bahasa Indonesia, hal ini dikarenakan dalam aplikasi RapidMiner sudah ada fasilitas kamus untuk mengubah akronim, dan stopword, tetapi masih terbatas pada bahasa Inggris, Cina, dan Arab, sedangkan untuk bahasa Indonesia masih belum tersedia. Dari hasil pre-processing dengan menggunakan Gata Framework, maka dataset akan dilakukan pre-processing lagi dengan menggunakan tools RapidMiner untuk membersihkan data agar lebih baik lagi hasilnya.

Gambar 2. Hasil pengolahan data review menggunakan Gata Framework (2020)

Tokenization: Regexp	Indonesian Stemming
mantap	mantap
oke	oke
mantep pak mantep	mantep pak mantep
saya suka sekali karena sangat membantu meetings jarak jauh apalagi disaat korona melanda di indonesia semua orang tidak boleh keluar rumah bisa belajar bareng	saya suka kali karena sangat bantu meetings jarak jauh apalagi saat korona landa di Indonesia semua orang tidak boleh keluar rumah bisa ajar bareng
saya baru akan mencoba untuk belajar pake zoom us terima kasih nggak bisa masuk app nya	saya baru akan coba untuk ajar pake zoom us terima kasih nggak bisa masuk app nya
terimakasih zoom	terimakasih zoom
di downlo nya lama banget ya dari tadi	di downlo nya lama banget ya dari tadi
bagus mas lur tapi belum banyak teman	bagus mas lur tapi belum banyak teman

Sumber: Hasil pengolahan data (2020)

Pada tahap ini adalah membuat model dengan menggunakan algoritma klasifikasi untuk dataset yang sudah melalui tahap pre-processing. Tahap ini menggunakan algoritma Naïve Bayes dengan menggunakan tools rapidminer versi 9.7 untuk mengolah dataset. Langkah dari proses ini adalah dataset akan diunggah ke dalam tools dengan menggunakan file excel yang akan di proses dengan algoritma Naïve Bayes untuk mendapatkan hasil awal dari masing-masing.

Jika hasil perhitungan awal dari penggunaan algoritma Naïve Bayes menunjukkan nilai accuracy dan kappa masih kurang bagus, maka langkah selanjutnya akan ditambahkan penggunaan feature selection untuk meningkatkan nilai accuracy maupun kappa.

Untuk pengujian menggunakan data review pengguna yang memberikan nilai skor penilaian bintang 1 dan bintang 5, sehingga hasil yang didapatkan seperti pada tabel dibawah ini.

Tabel 2. Perbandingan Nilai Accuracy dan AUC dengan Algoritma NB dan k-NN (metode kedua: 2 Kelas, Bintang 1 dan Bintang 5)

	Akurasi		AUC	
	Google Meet	Zoom Cloud Meeting	Google Meet	Zoom Cloud Meeting
Algoritma Naïve Bayes	66,74 %	72,42%	0,762	0,705

Sumber : Hasil penelitian (2020)

Dari hasil pengujian algoritma Naïve Bayes menghasilkan nilai Accuracy yang cukup baik yaitu diatas 60%. Jika pengujian kedua ini, menggunakan feature optimasi (SMOTE dan PSO) seperti yang dilakukan pada pengujian pertama, maka Algoritma K-Nearest Neighbor menghasilkan nilai Accuracy dan AUC mencapai 90% seperti pada tabel dibawah ini.

Tabel 3. Perbandingan Nilai Accuracy dan AUC dengan Algoritma NB dengan feature SMOTE Upsampling dan PSO (metode kedua: 2 Kelas, Bintang 1 dan Bintang 5)

	akurasi		AUC	
	Google Meet	Zoom Cloud Meeting	Google Meet	Zoom Cloud Meeting
<i>Algoritma Naïve Bayes+smote+PSO</i>	79,33%	85,76 %	0,852	0,961

Sumber: Hasil penelitian (2020)

Dapat disimpulkan bahwa feature selection SMOTE dan feature optimasi PSO sangat berpengaruh untuk meningkatkan nilai Accuracy, Kappa maupun AUC, namun sebaliknya jika tanpa menggunakan feature selection dan feature optimasi tidak mempengaruhi pada peningkatan nilai Accuracy, Kappa maupun AUC pada model algoritma Naïve Bayes (NB). Dari hasil pengujian dengan metode 2 kelas, maka dapat dihasilkan sebagai berikut:

1. Untuk pengujian dengan metode 2 kelas, maka algoritma Naïve Bayes dengan feature selection Synthetic Minority Over-Sampling Technique (SMOTE) dan feature optimasi Particle Swarm Optimization (PSO). Dimana untuk peningkatan nilai akurasi sebelum dan sesudah menggunakan feature selection SMOTE upsampling dan feature optimasi PSO dapat disampaikan pada tabel 4.
2. Aplikasi yang mendapatkan Akurasi yang tertinggi adalah aplikasi Zoom Cloud Meeting dengan pengujian awal 72,42 % dan setelah ditambahkan Smote dan PSO mendapatkan akurasi sebesar 85,76 %.

Tabel 4. Peningkatan Accuracy sebelum dan sesudah menggunakan SMOTE dan PSO (Metode Kedua – 2 Kelas)

Algoritma	Google Meet	Zoom Cloud Meeting
	Accuracy	
<i>Naïve Bayes</i>	66,74 %	72,42 %
<i>Naïve Bayes+SMOTE+PSO</i>	79,33 %	85,76 %
Progres Kenaikan Accuracy	12,59 %	13,34 %

Sumber: Hasil penelitian (2020)

Dari tabel diatas peningkatan nilai akurasi dengan menggunakan algoritma Naïve Bayes sebelum dan dan sesudah menggunakan feature selection SMOTE Upsampling dan feature optimasi PSO untuk analisis review pengguna Google Meet

mencapai 12,59 %, sedangkan untuk review pengguna Zoom Cloud Meeting kenaikan prosentasenya mencapai 13,34 %.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat ditarik beberapa kesimpulan bahwa dari hasil perhitungan algoritma klasifikasi Naïve Bayes dengan feature selection Synthetic Minority Over-Sampling Technique (SMOTE) dan feature optimasi Particle Swarm Optimization (PSO) dalam mengolah data review pengguna Google Meet dan Zoom Cloud Meeting yang diambil dari google play sebanyak 714 data, menunjukkan bahwa adanya peningkatan nilai Accuracy, Kappa maupun AUC setelah menggunakan feature selection SMOTE Upsampling dan feature optimasi PSO. Hasil kesimpulannya sebagai berikut:

1. Untuk pengujian dengan metode 2 kelas, maka algoritma Naïve Bayes dengan feature selection Synthetic Minority Over-Sampling Technique (SMOTE) dan feature optimasi Particle Swarm Optimization (PSO). Dimana untuk peningkatan nilai akurasi sebelum dan sesudah menggunakan feature selection SMOTE upsampling dan feature optimasi PSO dapat disampaikan pada tabel 4.
2. Aplikasi yang mendapatkan Akurasi yang tertinggi adalah aplikasi Zoom Cloud Meeting dengan pengujian awal 72,42 % dan setelah ditambahkan Smote dan PSO mendapatkan akurasi sebesar 85,76 %.

REFERENSI

- Abdillah, G., Putra, F. A., & Renaldi, F. (2016). Penerapan Data Mining Pemakaian Air Pelanggan Untuk Menentukan Klasifikasi Potensi Pemakaian Air Pelanggan Baru Di Pdam Tirta Raharja Menggunakan Algoritma K-Means. *Sentika 2016, 2016*(Sentika), 18–19.
- Buntoro, G. A. (2017). Analisis Sentimen Calon Gubernur DKI Jakarta 2017 Di Twitter. *Integer Journal Maret, 1*(1), 32–41. https://www.researchgate.net/profile/Ghulam_Buntoro/publication/316617194_Analisis_Sentimen_Calon_Gubernur_DKI_Jakarta_2017_Di_Twitter/links/5907eee44585152d2e9ff992/Analisis-Sentimen-Calon-Gubernur-DKI-Jakarta-2017-Di-Twitter.pdf
- Darmawan, A., & Syamsiah, S. (2016). Analisa Model Support Vector Machine Textmining pada Komentar Positif dan Negatif untuk Review Perbandingan Whatsapp Vs BBM. *STRING (Satuan Tulisan Riset Dan Inovasi Teknologi), 1*(1), 74–82. <https://doi.org/10.30998/string.v1i1.971>
- Gunadi, G., & Sensuse, D. I. (2012). Penerapan Metode Data Mining Market Basket Analysis

- Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) : Studi Kasus Percetakan PT. Gramedia. *Telematika*, 4(1), 118–132.
- Gunawan, F., Fauzi, M. A., & Adikara, P. P. (2017). Analisis Sentimen Pada Ulasan Aplikasi Mobile Menggunakan Naive Bayes dan Normalisasi Kata Berbasis Levenshtein Distance (Studi Kasus Aplikasi BCA Mobile). *Systemic: Information System and Informatics Journal*, 3(2), 1–6. <https://doi.org/10.29080/systemic.v3i2.234>
- Hanggara, S., Akhriza, T. M., & Husni, M. (2017). Aplikasi Web Untuk Analisis Sentimen Pada Opini Produk Dengan Metode Naive Bayes Classifier. *Seminar Nasional Inovasi Dan Aplikasi Teknologi Di Industri*, 4(2), A33.1-A33.6.
- Hartanto. (2017). Text Mining Dan Sentimen Analisis Twitter Pada Gerakan Lgbt. *Intuisi : Jurnal Psikologi Ilmiah*, 9(1), 18–25.
- Hulu, A. M. F., & Lhaksana, K. M. (2019). Analisis Sentimen pada Twitter Menggunakan Metode Support Vector Machine. *Seminar Nasional Teknologi Fakultas Teknik Universitas Krisnadwipayana*, 1(1), 739–742.
- Minarni, & Irawan, P. (2019). Implementasi Metode Naive Bayes Untuk Diagnosa Penyakit Lambung. *Jurnal TEKNOIF*, 9986(September), 115–123. <https://doi.org/10.21063/JTIF.2019.V7.2>.
- Pertiwi, M. W. (2019). Analisis sentimen opini publik mengenai sarana dan transportasi mudik tahun 2019 pada twitter menggunakan algoritma naïve bayes, neural network, KNN dan SVM. *Inti Nusa Mandiri*, 14(1), 27–32.
- Raco, J. R. (2010). *Metode Penelitian Kualitatif: Jenis, Karakteristik dan Keunggulannya*.
- Ratniasih, N. L., Sudarma, M., & Gunantara, N. (2017). Penerapan Text Mining Dalam Spam Filtering Untuk Aplikasi Chat. *Majalah Ilmiah Teknologi Elektro*, 16(3), 13. <https://doi.org/10.24843/mite.2017.v16i03p03>
- Rofiqoh, U., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler Indonesia Pada Twitter Dengan Metode Support Vector Machine dan Lexion Based Feature. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 1(12), 1725–1732. <http://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/628>
- Salam, A., Zeniarja, J., & Khasanah, R. S. U. (2018). Analisis Sentimen Data Komentar Sosial Media Facebook Dengan K-Nearest Neighbor (Studi Kasus Pada Akun Jasa Ekspedisi Barang J&T Ekpress Indonesia). *Prosiding SINTAK*, 480–486.
- Saritas, M. M., & Yasar, A. (2019). Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification. *Faraday Discuss.*, 2(7), 88–91. <https://doi.org/10.1039/b000000x>