

Analisis Sentimen Ulasan Aplikasi Gopay Menggunakan *Naive Bayes* dengan Teknik *Oversampling*

Ita Yulianti¹, Lis Saumi Ramdhani², Tya Septiani Nurfauzia Koeswara³

^{1,2}Sistem Informasi Akuntansi Kampus Kota Sukabumi, Universitas Bina Sarana Informatika
Sukabumi, Indonesia

³Sistem Informasi Kampus Kota Sukabumi, Universitas Bina Sarana Informatika
Sukabumi, Indonesia

e-mail: ita.iyi@bsi.ac.id¹, lis.lud@bsi.ac.id², tya.tsf@bsi.ac.id³

ABSTRAK

Aplikasi Gopay merupakan salah satu dari aplikasi fintech yang paling banyak digunakan dibandingkan dengan platform lainnya dan tercatat memiliki ulasan sebanyak 207.000 di kolom komentar pada Play Store. Dengan memanfaatkan data tersebut, penelitian ini bertujuan untuk mengekstraksi dan menganalisis secara komputasional menjadi sentimen positif dan negatif menggunakan algoritma *Naive Bayes* yang dikombinasikan dengan teknik *oversampling*. Adapun untuk dataset yang digunakan bersumber dari Google Play Store. Hasil penelitian menunjukkan bahwa pemodelan yang diterapkan mampu menghasilkan nilai akurasi yang mencapai 92,89%. Hal ini menjadi temuan penting dikarenakan adanya peningkatan akurasi yang signifikan jika dibandingkan dengan pemodelan *Naive Bayes* tanpa menggunakan *oversampling* yang hanya 89,19% saja. Dengan demikian, penerapan teknik *oversampling* dalam model *Naive Bayes* ini dapat menjadi solusi efektif dalam mengatasi ketidakseimbangan kelas pada analisis sentimen karena mampu meningkatkan performa model klasifikasi.

Kata Kunci: analisis sentimen, gopay, *naive bayes*, *oversampling*, play store

ABSTRACTS

The Gopay application is one of the most widely used fintech applications compared to other platforms and has recorded 207,000 reviews in the comments column on the Play Store. By utilizing this data, this research aims to extract and computationally analyze positive and negative sentiment using the Naive Bayes algorithm combined with oversampling techniques. The dataset used is sourced from the Google Play Store. The research results showed that the modeling applied was able to produce an accuracy value of 92.89%. This is an important finding because there is a significant increase in accuracy when compared to Naive Bayes modeling without using oversampling which is only 89.19%. Thus, the application of oversampling techniques in the Naive Bayes model can be an effective solution in overcoming class imbalance in sentiment analysis because it can improve the performance of the classification model.

Keywords: gopay, *naive bayes*, *oversampling*, play store, sentiment analysis

1. PENDAHULUAN

Kebutuhan akan pemrosesan data yang mudah dan cepat saat ini mendorong berbagai perusahaan maupun instansi untuk menciptakan

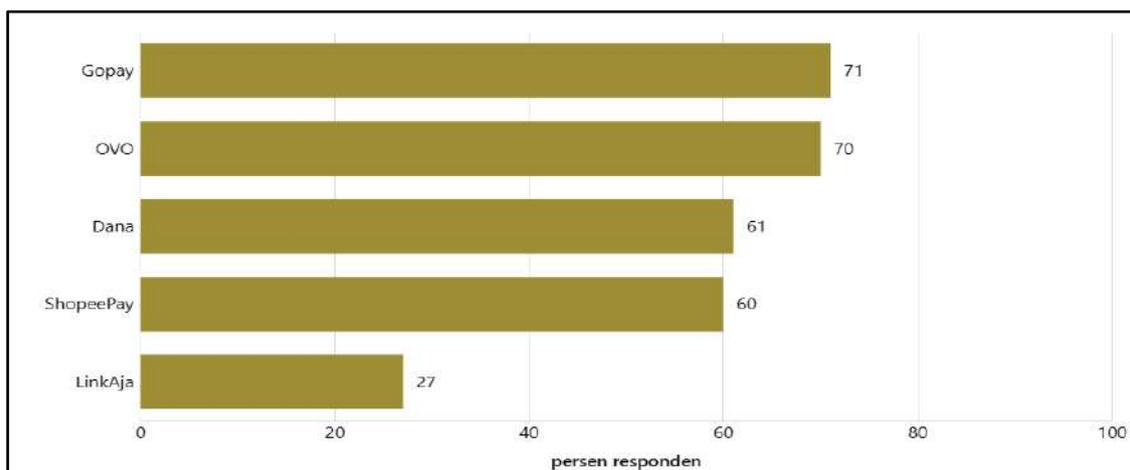
inovasi baru pada seluruh lini kehidupan (Tanggraeni & Sitokdana, 2022). Peranan teknologi memiliki dampak yang signifikan dalam memicu hadirnya inovasi-inovasi tersebut (Riskawati et al., 2024) termasuk dalam bidang



pendidikan, transportasi, hingga menyentuh sektor keuangan. Salah satu inovasi yang muncul sebagai katalisator utama dalam transformasi sektor keuangan adalah *Financial Technology* atau biasa dikenal dengan istilah Fintech (Tri Dewi Septiani et al., 2023). FinTech mampu menggeser peran dari lembaga keuangan formal seperti bank menjadi sistem pembayaran baru (Ghozali et al., 2022) yang lebih mudah, praktis dan aman dalam mengelola layanan keuangan secara umum (Ansori et al., 2024). Ada berbagai fitur yang ditawarkan pada *platform* fintech, diantaranya proses pembayaran, transfer, jual-beli saham, peminjaman modal, dan masih banyak lagi yang lainnya (Basalamah et al., 2022).

Pada Festival Ekonomi Keuangan Digital Indonesia (FEKDI) tahun 2022 menunjukkan bahwa ada 5 platform fintech yang mendominasi masyarakat Indonesia dalam kategori pembayaran (*e-wallet*) yaitu Gopay, DANA, OVO, ShopeePay, dan LinkAja (Maharani et al., 2022).

Pengguna *e-wallet* tersebut mayoritas melakukan transaksi untuk berbagai kebutuhan seperti berbelanja *online*, pembelian pulsa/paket data, tagihan, transportasi dan pembelian makanan (Ningri et al., 2023). Gambar 1 menunjukkan grafik yang menampilkan platform *e-wallet* yang paling sering digunakan oleh masyarakat Indonesia. Gambar 1. merupakan laporan *E-Wallet Industry Outlook 2023* dari Insight Asia yang melakukan survey pada September 2022 yang menunjukkan bahwa Gopay merupakan platform *e-wallet* yang paling banyak digunakan dibanding platform lainnya dengan proporsi pengguna sebanyak 71%. Jutaan orang telah memanfaatkan *e-wallet* tersebut untuk mengelola transaksi keuangan secara digital yang memberikan kemudahan dan menghindari terjadinya kerumitan transaksi fisik yang terkadang dirasa lebih merepotkan (Rohmah & Kurnianingsih, 2023).



Sumber: Insight Asia (Ahdiat, 2023)

Gambar 1. Hasil Survei Penggunaan E-Wallet di Indonesia pada September 2022

Kelebihan aplikasi Gopay salah satunya adalah memiliki fitur simpanan yang dirancang dengan *reminder* yang bertujuan untuk menumbuhkan kedisiplinan dalam menabung bagi para pengguna, serta terdapat fitur laporan pengeluaran yang dapat digunakan untuk memantau dan mengontrol pengeluaran sehingga dapat membantu dalam mencapai tujuan keuangan. Saat ini, Aplikasi Gopay telah diunduh oleh 10 juta lebih pengguna Play Store dengan rating 4.6 dan tercatat memiliki ulasan sebanyak 207.000 di kolom komentar. Hal ini bisa menjadi sumber informasi yang berharga bagi perusahaan dalam mengukur kualitas layanan yang diberikan dan sebagai referensi bagi

pengguna lain untuk mempertimbangkan menggunakan aplikasi tersebut (Irnawati & Solecha, 2022). Akan tetapi, banyaknya komentar yang tersedia, tidak menjamin keakuratan dari pengalaman pengguna, sehingga terkadang ada ulasan positif maupun negatif yang tidak relevan dengan rating yang diberikan (Larasati et al., 2022). Maka dari itu, agar dapat menganalisis opini masyarakat dalam penggunaan aplikasi tersebut, maka pada penelitian ini akan dilakukan analisis sentimen berdasarkan ulasan pengguna yang bersumber dari Play Store.

Analisis sentimen merupakan suatu teknik pemrosesan data teks yang menggunakan

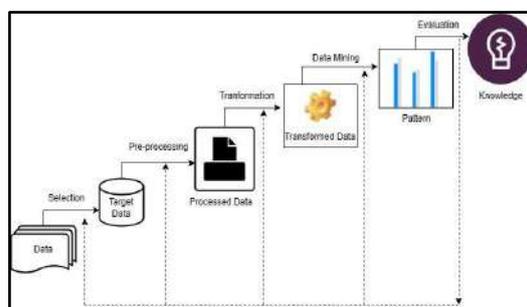
pembobotan pada suatu kalimat atau pasangan kata untuk mendeteksi opini dan mengklasifikannya menjadi kelas positif atau negatif (Indarwati & Februariyanti, 2023). Beberapa contoh dari penelitian sebelumnya yang telah membahas mengenai studi tersebut, diantaranya; Penelitian (Mahendrajaya et al., 2019) yang memanfaatkan *tools* WEKA untuk menganalisis pengguna Gopay melalui opini di twitter menggunakan metode *Lexion Based* dan *Support Vector Machine* sehingga diperoleh akurasi yang cukup baik senilai 89,17%. Kemudian, penelitian dari (Ghozali et al., 2022) yang melakukan analisis sentimen terhadap pinjaman *online* menggunakan metode *Naive Bayes* yang mendapatkan tingkat akurasi sebesar 80%. Di sisi lain, ada juga penelitian yang membandingkan antara algoritma *Naive Bayes Classifier* dan *K-Nearest Neighbor* dalam menganalisis ulasan aplikasi PLN *Mobile* dan hasilnya menyatakan bahwa pemodelan dengan algoritma NBC lebih baik dibandingkan KNN dengan perolehan akurasi sebesar 77,69% (Syafrizal et al., 2023). Selain itu, penelitian lainnya ada dari (Riskawati et al., 2024) yang mengkaji efektivitas metode *Naive Bayes* dalam analisis sentimen aplikasi Gopay berdasarkan data ulasan Play Store dari Agustus s/d November 2023 dengan hasil temuan akurasi bernilai 85%.

Merujuk pada penelitian-penelitian tersebut, dapat disimpulkan bahwa algoritma *Naive Bayes* telah terbukti keberfamaannya untuk memproses data dalam berbagai konteks dan aplikasi. Algoritma ini seringkali digunakan dalam berbagai studi seperti klasifikasi, menyaring spam, analisis sentimen, dan lainnya (Syafrizal et al., 2023). Meskipun penelitian sebelumnya telah berhasil dilakukan, namun sebagian besar prosesnya cenderung mengabaikan penggunaan bahasa *slang* yang sering muncul dalam ulasan, sehingga menyebabkan model sulit memahami konteks sentimen yang tepat. Oleh karena itu, tujuan dari penelitian ini ialah menerapkan algoritma *Naive Bayes* pada analisis sentimen pengguna aplikasi Gopay dengan menggunakan pendekatan yang lebih adaptif terhadap bahasa informal melalui dataset *terupdate* yang diintergrasikan dengan teknik *oversampling* untuk mengoptimalkan akurasi klasifikasi pada dataset yang tidak seimbang. Dengan adanya penelitian ini diharapkan dapat memberikan kontribusi antara lain dapat membantu pengembang sistem dalam mengoptimalkan layanan yang diberikan dengan

menganalisis respon pengguna agar fitur yang dikembangkan sesuai dengan preferensi dan kebutuhan pengguna. Selain itu, dari perspektif pengguna, penelitian ini juga bisa menjadi salah satu informasi penunjang bagi calon pengguna sebelum memutuskan untuk bergabung atau tidak pada aplikasi tersebut.

2. METODE PENELITIAN

Metode penelitian yang diterapkan pada penelitian ini yaitu dengan menggunakan prinsip dan teknik dari KDD (*Knowledge Discovery in Databases*) yang bertujuan untuk menghasilkan informasi dan pola dari data yang telah dipilih sebelumnya (Khoirul Insan et al., 2023). Adapun alur dari teknik ini meliputi *Selection*, *Preprocessing*, *Transformation*, *Data Mining*, dan *Evaluation*, yang dapat dilihat pada gambar 2.



Gambar 2. Tahapan Penelitian

Tahapan *selection* melibatkan penentuan dataset yang akan digunakan. Pada penelitian ini dataset yang dipilih bersumber dari ulasan Google Play Store yang diperoleh menggunakan API resmi Google Play Developer yaitu Google Play Scraper.

Pada penelitian ini, tahapan *preprocessing* diawali dengan melakukan pelabelan data. Dataset kemudian diproses kembali kedalam serangkaian tahap pemrosesan meliputi *Case Folding*, *Stopword Removal*, *Tokenizing*, dan *Stemming* yang dilakukan secara berurutan sehingga diperoleh dataset yang lebih bersih dengan tingkat *noise* lebih rendah dibandingkan sebelumnya. Setelah itu, dilanjutkan dengan pembuatan *word cloud* dan penerapan teknik *oversampling* jika diperlukan untuk mengatasi *imbalance class* pada dataset. Tujuan utama dijalankannya tahapan *preprocessing* ini adalah untuk membersihkan dan mempersiapkan data serta memastikan data yang akan dianalisis dalam kondisi optimal.

Tahapan *transformation* berarti mengkonversi data kedalam format yang sesuai

untuk dianalisis. Pada penelitian ini, ada dua metode transformasi yang digunakan yaitu metode TF-IDF (*Term Frequency-Inverse Document Frequency*) dan *Term Frequency Vectorizer* (TFV). Kedua metode ini sama-sama berkontribusi dalam memperkaya dimensi data, sehingga model yang dibangun dapat mengenali pola-pola yang bermakna dalam ulasan.

Tahapan *data mining* melibatkan penerapan algoritma yang akan digunakan untuk mengidentifikasi ulasan kedalam dua kelas target yaitu sentimen positif dan negatif. Pada penelitian ini, proses pemahaman pola dan prediksi sentimen tersebut diaplikasikan menggunakan algoritma *Multinomial Naive Bayes*.

Tahapan evaluasi dilaksanakan untuk mengetahui kinerja dari algoritma yang diterapkan. Pada tahap terakhir ini, diperoleh hasil *Classification Report* dan bagan *Confusion Matrix* yang menunjukkan informasi berupa nilai akurasi, presisi, *recall*, dan *F1 Score*.

3. HASIL DAN PEMBAHASAN

Pada penelitian ini, analisis data dilakukan menggunakan bahasa pemrograman python yang dieksekusi melalui *tools* Google Colaboratory. Dengan bantuan *tools* tersebut memungkinkan proses pengolahan dan pembangunan model serta pembuatan visualisasi menjadi lebih optimal.

3.1. Selection

Pemilihan dan pengumpulan dataset dalam penelitian ini dilakukan dengan scraping data ulasan aplikasi Gopay dari Google Play Store. Dari hasil scraping diperoleh informasi meliputi username, rating, tanggal dan isi ulasan, yang kemudian diexport kedalam format file csv. Sebanyak 796 ulasan berhasil dikumpulkan berdasarkan data yang paling relevan dengan rentang waktu dari bulan Januari s/d Agustus 2024.

3.2. Preprocessing

Tahapan *preprocessing* terdiri dari berbagai serangkaian proses yaitu pelabelan, case folding, stopward removal, tokenizing, stemming, visualisasi word cloud, dan oversampling.

3.2.1. Pelabelan

Proses ini dimulai dengan menghilangkan atribut yang tidak berpengaruh penting pada *preprocessing* data. Dari semua atribut yang diperoleh dalam tahapan *selection*, hanya dua atribut yang digunakan yaitu atribut rating dan isi ulasan. Setelah itu, dilakukan proses pelabelan dengan memberikan nilai label sentimen positif atau negatif yang ditentukan berdasarkan rating pada ulasan. Pada proses ini, atribut label ditambahkan sebagai kelas target pada dataset seperti yang terlihat pada gambar 3.

	ulasan	rating	label
0	Jujur kecewa sih sama customer service nya,lap...	1	Negatif
1	Cara pakainya mudah cuma entah kenapa Akun say...	5	Positif
2	Tiba tiba akun GoPay Plus saya jadi Regular Ke...	4	Positif
3	Aplikasi tidak jelas saya tidak bisa transfer ...	1	Negatif
4	bagus , cuma kelamaan transaksi gara gara ga l...	5	Positif
5	the Best pokoknya UI nya keren mudah dipahami ...	5	Positif
6	Gangguan terus, buang waktu, pemrosesan lama b...	1	Negatif
7	Kerja lebih dari 3 tahun di go car mau pake pa...	1	Negatif
8	kenapa ya, kok nggk bisa upgrade ke plus? pdhl...	1	Negatif
9	Pelayanan sangat buruk, hanya karena mengganti...	1	Negatif

Gambar 3. Tampilan Sampel Hasil Pelabelan

Langkah selanjutnya adalah memastikan bahwa tidak ada data yang *null* dalam dataset dengan menggunakan metode *isnull()* dan metode *dropna()*. Setelah proses ini dilakukan, dataset yang awalnya berjumlah 796 ulasan berkurang menjadi 738 ulasan, menunjukkan bahwa 58 ulasan teridentifikasi *null* dan telah dihapus seperti yang terlihat pada gambar 4.

```
Jumlah data yang dihapus: 58
Jumlah kelas target: 738
label
Negatif    470
Positif     268
Name: count, dtype: int64
```

Gambar 4. Hasil Data Setelah Pembersihan *Null*

3.2.2. Case Folding

Pada tahapan ini seluruh teks pada dataset dikonversi menjadi huruf kecil. Selain itu, dijalankan juga proses pembersihan teks dari karakter yang tidak relevan seperti emoji dan tanda baca dengan menggunakan teknik *regex* dengan hasil yang terlihat pada tabel 1.

Tabel 1. Hasil *Case Folding*

Sebelum	Sesudah
Entah kenapa gopay kalau mau jadikan metode pembayaran pasti loading nya lama banget sampai berkali kali nge stuck di loading terus 🙄🙄	entah kenapa gopay kalau mau jadikan metode pembayaran pasti loading nya lama banget sampai berkali kali nge stuck di loading terus
Aplikasinya sangat membantu buat transaksi kapan saja di mana saja dengan respon cepat, sukses terus buat gopay 👍👍	aplikasinya sangat membantu buat transaksi kapan saja di mana saja dengan respon cepat sukses terus buat gopay
Proses pembayarannya cepat, Aku suka banget top up di Gopay. nyesel baru make sekarang tolong jangan langsung ke log out sendiri yaa selama aku ga buka 🙄🙄🙄	proses pembayarannya cepat aku suka banget top up di gopay nyesel baru make sekarang tolong jangan langsung ke log out sendiri yaa selama aku ga buka

3.2.3. Stopword Removal

Langkah berikutnya yaitu *Stopword Removal* yang bertujuan untuk menghilangkan kata yang tidak memiliki makna penting dengan tujuan analisis. Selain menerapkan daftar *stopwords* yang tersedia di python, dalam penelitian ini juga ditambahkan *custom stopwords* yang dikembangkan berdasarkan analisis mendalam terhadap dataset yang digunakan. *Custom stopwords* dalam penelitian ini berisi berbagai istilah *slang* seperti “nih”, “banget”, “dong” dan yang lainnya. Hasil dari proses ini dapat dilihat pada tabel 2.

Tabel 2. Hasil *Stopword Removal*

Sebelum	Sesudah
[entah, kenapa, gopay, kalau, mau, jadikan, metode, pembayaran, pasti, loading, nya, lama, banget, sampai, berkali, kali, nge, stuck, di, loading, terus]	gopay jadikan metode pembayaran loading stuck loading
[aplikasinya, sangat, membantu, buat, transaksi, kapan, saja, di, mana, saja, dengan, respon, cepat, sukses, terus, buat, gopay]	aplikasinya membantu transaksi respon cepat sukses gopay

3.2.4. Tokenizing

Proses *Tokenizing* pada penelitian ini memanfaatkan library NLTK yaitu *resource* ‘punkt’ yang membantu membagi setiap record pada dataset menjadi bagian yang lebih kecil (token) dan lebih mudah dikelola seperti yang terlihat pada tabel 3.

Tabel 3. Hasil *Tokenizing*

Sebelum	Sesudah
gopay jadikan metode pembayaran loading stuck loading aplikasinya membantu transaksi respon cepat sukses gopay	[gopay, jadikan, metode, pembayaran, loading, stuck, loading] [aplikasinya, membantu, transaksi, respon, cepat, sukses, gopay]

3.2.5. Stemming

Pada proses *Stemming*, setiap kata pada dataset dikonversi kedalam bentuk dasarnya. Dalam penelitian ini, proses tersebut dilakukan dengan menerapkan pustaka sastra yang dapat digunakan khusus untuk ulasan yang berbahasa Indonesia dengan hasil yang terlihat pada tabel 4.

Tabel 4. Hasil *Stemming*

Sebelum	Sesudah
[gopay, jadikan, metode, pembayaran, loading, stuck, loading] [aplikasinya, membantu, transaksi, respon, cepat, sukses, gopay]	gopay jadi metode bayar loading stuck loading aplikasi bantu transaksi respon cepat sukses gopay

3.2.6. Visualisasi Word Cloud

Word Cloud berisi frekuensi dari semua kata yang ada pada dataset. Visualisasi dari *word cloud* ini memungkinkan untuk mengidentifikasi kata kunci dan tema utama dari dataset secara visual serta memberikan temuan yang relevan untuk analisa lebih lanjut. Tampak pada gambar 5, visualisasi *word cloud* yang menunjukkan preferensi positif dan negatif terhadap semua fitur yang ada pada aplikasi Gopay.

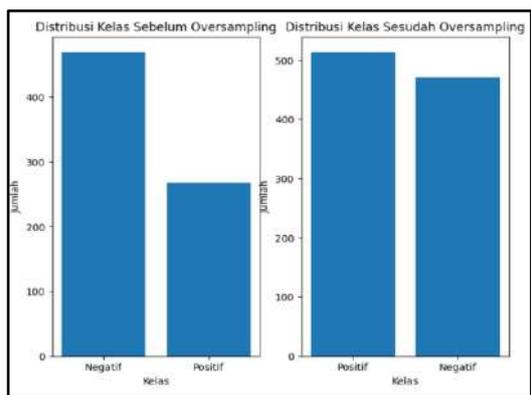
3.2.7. Oversampling

Langkah berikutnya adalah melakukan pengecekan *imbalance class*. Tahapan ini penting dilakukan supaya model yang akan dikembangkan tidak bias terhadap kelas tertentu. Pada dataset penelitian ini (Lihat Gambar 4.) menunjukkan terjadinya *imbalance*

class, dimana proporsi nilai kelas “Negatif” lebih dominan dibandingkan kelas “Positif”. Hal tersebut dapat berdampak terhadap kinerja model yang mungkin akan cenderung untuk memprediksi kelas “Negatif”. Oleh karena itu, teknik *oversampling* digunakan untuk menangani masalah tersebut. Cara kerja dari teknik ini yaitu dengan mengidentifikasi kelas minoritas dan *resample* untuk menambah jumlah data hingga mencapai proporsi sesuai dengan kebutuhan analisis. Gambar 6 menampilkan hasil visualisasi yang menunjukkan perubahan yang terjadi setelah proses *oversampling*.



Gambar 5. Visualisasi Word Cloud



Gambar 6. Perbandingan Sebelum dan Sesudah Proses *Oversampling*

3.3. Transformation

Tahapan adalah transformasi, dimana tahapan ini perlu dijalankan supaya dataset bisa diterapkan pemodelan menggunakan algoritma machine learning. Tahapan ini dimulai dengan menerapkan metode TF-IDF agar dataset yang

awalnya berupa teks berubah menjadi vektor numerik berbasis TF-IDF. Kemudian, metode TFV digunakan untuk mengkalkulasikan frekuensi kata pada dataset sehingga dihasilkan matriks fitur berbasis *count*. Dengan mengkombinasikan kedua metode tersebut, maka dataset telah menjadi format numerik dan siap diproses lebih lanjut terutama untuk evaluasi model.

3.4. Data Mining

Pada tahapan ini algoritma *Multinomial Naive Bayes* digunakan untuk pembangunan model dalam prediksi analisis sentimen aplikasi Gopay. Algoritma ini dipilih karena kemampuannya dalam mengolah data teks yang telah ditransformasi menggunakan TF-IDF dan TFV. Pada proses *data mining*, dataset dibagi menjadi dua bagian yang terdiri dari data *training* dan data *testing* dengan rasio 80:20. Dari hasil penelitian ini diperoleh bahwa model analisis sentimen yang dibangun mencapai akurasi sebesar 92,89%. Hal ini mengindikasikan bahwa pemodelan tersebut mampu mengklasifikasikan sentimen pengguna Gopay dengan efektif.

MultinomialNB Accuracy: 0.9289340101522843				
MultinomialNB Precision: 0.9411764705882353				
MultinomialNB Recall: 0.9230769230769231				
MultinomialNB f1_score: 0.9320388349514563				
=====				
	precision	recall	f1-score	support
Negatif	0.94	0.92	0.93	104
Positif	0.92	0.94	0.93	93
accuracy			0.93	197
macro avg	0.93	0.93	0.93	197
weighted avg	0.93	0.93	0.93	197

Gambar 7. Hasil Pemodelan

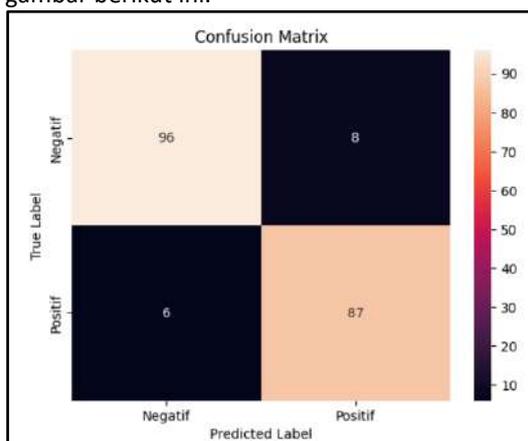
3.5. Evaluation

Untuk memastikan kinerja model yang dibangun telah optimal, maka perlu dilakukan evaluasi dengan beberapa rangkaian proses. Pertama, melakukan perbandingan dengan pemodelan yang sama namun tanpa menerapkan teknik *oversampling*. Hasil perbandingan ini dapat dilihat pada Tabel 5.

Tabel 5. Hasil Evaluasi Matriks

Matriks Evaluasi	Pemodelan NB dengan <i>Oversampling</i>	Pemodelan NB
Akurasi	92,89%	89,19%
Presisi	94,12%	87,96%
Recall	92,31%	96,94%
F1 Score	93,20%	92,23%

Berdasarkan Tabel 5. dapat diketahui bahwa pemodelan Naive Bayes dengan *oversampling* lebih baik dibandingkan tanpa *oversampling* terutama dalam hal akurasi. Hal ini dibuktikan dengan terjadinya peningkatan yang signifikan sebesar 3,7% yang menunjukkan bahwa model tersebut lebih akurat dalam mengklasifikasikan data sentimen secara keseluruhan. Selain mengukur dengan evaluasi matriks, dilakukan juga analisis menggunakan *confusion matrix* yang hasilnya dapat dilihat pada gambar berikut ini.



Gambar 8. Hasil Evaluasi dengan *Confusion Matrix*

Gambar 8 merupakan visualisasi dari distribusi prediksi model terhadap kelas sentimen positif dan negatif. Adapun faktor kunci yang berkontribusi terhadap peningkatan performa ini diantaranya penggunaan *custom stopword* yang membantu model menjadi lebih berfokus pada kata-kata yang berpengaruh terhadap sentimen dan implementasi dari teknik *oversampling* untuk mengatasi masalah ketidakseimbangan kelas. Melalui kombinasi dari kedua teknik tersebut, kemampuan model menjadi meningkat sehingga dapat mencapai hasil yang lebih baik jika dibandingkan dengan penelitian sebelumnya yang mungkin belum menerapkan pendekatan serupa.

4. KESIMPULAN

Penelitian ini dilakukan untuk menganalisis sentimen preferensi pengguna terhadap semua fitur yang tersedia dalam aplikasi Gopay berdasarkan ulasan yang diambil melalui teknik *scraping* pada Google Play Store. Adapun model klasifikasi yang diterapkan yaitu menggunakan algoritma *Multinomial Naive Bayes* dengan teknik *oversampling*. Hasil penelitian menunjukkan bahwa pemodelan

tersebut dinilai efektif dalam memprediksi dataset yang diuji. Hal ini dapat dibuktikan salah satunya melalui tingkat akurasi model yang dihasilkan yakni dari 89,19% meningkat sebanyak 3,7% sehingga menjadi 92,89%. Dengan demikian, penerapan teknik *oversampling* dalam model Naive Bayes ini dinyatakan mampu meningkatkan performa model klasifikasi secara signifikan.

Saran penelitian selanjutnya, diharapkan dapat menggunakan dataset lebih besar dan mengembangkan kualitas *preprocessing* data, serta mengeksplorasinya dengan algoritma lain seperti Support Vector Machine (SVM) atau K-Nearest Neighbors (KNN) agar dapat memberikan hasil yang lebih optimal.

5. REFERENSI

- Ahdiat, A. (2023). *Survei Pengguna Dompnet Digital: Gopay dan OVO Bersaing Ketat*. Databoks. <https://databoks.katadata.co.id/datapublish/2023/07/25/survei-pengguna-dompnet-digital-gopay-dan-ovo-bersaing-ketat>
- Ansori, M. A. Z., Wahyudin, M. W., Nurbaet, N. I., Isagozi, M. R., Diva, S. A., Zahra, N. A., Nur, G., Yusuf, M., & Tabroni, M. (2024). Analisis Literasi Keuangan Penggunaan Fintech Payment Bagi Pelaku Usaha UMKM. *Economic Reviews Journal*, 3(1), 210–225. <https://doi.org/10.56709/mrj.v3i1.130>
- Basalamah, R., Nurdin, N., Haekal, A., Abdul, J., & Noval, N. (2022). Risiko Terhadap Minat Menggunakan Financial Technology (Fintech) Gopay. *Jurnal Ilmu Ekonomi Dan Bisnis Islam*, 4(1), 57–71.
- Ghozali, M. I., Sugiharto, W. H., & Iskandar, A. F. (2022). Analisis Sentimen Pinjaman Online Di Media Sosial Twitter Menggunakan Metode Naive Bayes. *KLIK: Kajian Ilmiah Informatika Dan Komputer*, 33(1), 1–12. <https://doi.org/10.30865/klik.v3i6.936>
- Indarwati, K. D., & Februariyanti, H. (2023). Analisis Sentimen Terhadap Kualitas Pelayanan Aplikasi Gojek Menggunakan Metode Naive Bayes Classifier. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 10(1). <https://doi.org/10.35957/jatisi.v10i1.2643>
- Irnowati, O., & Solecha, K. (2022). Analisis Sentimen Ulasan Aplikasi Flip

- Menggunakan Naive Bayes dengan Seleksi Fitur PSO. *Jurnal Ilmiah Intech*, 4(2), 189–199.
- Khoirul Insan, M. K., Hayati, U., & Nurdiawan, O. (2023). Analisis Sentimen Aplikasi Brimo Pada Ulasan Pengguna Di Google Play Menggunakan Algoritma Naive Bayes. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 478–483. <https://doi.org/10.36040/jati.v7i1.6373>
- Larasati, F. A., Ratnawati, D. E., & Hanggara, B. T. (2022). Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 6(9), 4305–4313. <http://j-ptiik.ub.ac.id>
- Maharani, F. M. D., Hananto, A. L., Hilabi, S. S., Apriani, F. N., Hananto, A., & Huda, B. (2022). Perbandingan Metode Klasifikasi Sentimen Analisis Penggunaan E-Wallet Menggunakan Algoritma Naïve Bayes dan K-Nearest Neighbor. *Metik Jurnal*, 6(2), 97–103. <https://doi.org/10.47002/metik.v6i2.372>
- Mahendrajaya, R., Buntoro, G. A., & Setyawan, M. B. (2019). Analisis Sentimen Pengguna Gopay Menggunakan Metode Lexicon Based Dan Support Vector Machine. *Komputek*, 3(2), 52. <https://doi.org/10.24269/jkt.v3i2.270>
- Ningri, L. J., Hamidi, M., & Adrianto, F. (2023). Sentiment Analysis Against Digital Payment “GoPay”, “OVO”, “DANA”, and “ShopeePay” Using Naïve Bayes Classifier Algorithm. *Indonesian Journal of Economics and Management*, 3(2), 322–336. <https://doi.org/10.35313/ijem.v3i2.4765>
- Riskawati, Fatihanursari, Iin, & Rinaldi, A. R. (2024). Penerapan Metode Naïve Bayes Classifier Pada Analisis Sentimen Aplikasi Gopay. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(1), 346–353.
- Rohmah, N. A., & Kurnianingsih, H. (2023). Faktor - Faktor Yang Mempengaruhi Pengguna Aplikasi Gopay. *JEBI: Jurnal Ekonomi Dan Bisnis*, 1(6), 890–904.
- Syafrizal, S., Afdal, M., & Novita, R. (2023). Analisis Sentimen Ulasan Aplikasi PLN Mobile Menggunakan Algoritma Naïve Bayes Classifier dan K-Nearest Neighbor. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(1), 10–19. <https://doi.org/10.57152/malcom.v4i1.983>
- Tanggraeni, A. I., & Sitokdana, M. N. N. (2022). Analisis Sentimen Aplikasi E-Government pada Google Play Menggunakan Algoritma Naïve Bayes. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(2), 785–795. <https://doi.org/10.35957/jatisi.v9i2.1835>
- Tri Dewi Septiani, A., Prayogo Kuncoro, A., Subarkah, P., & Riyanto. (2023). Perbandingan Kinerja Metode Naïve Bayes Classifier dan K-Nearest Neighbor pada Analisis Sentimen Ulasan Mobile Banking Jenius. *Jurnal Krisnadana*, 3, 67–77. <https://ejournal.sidyanusa.org/index.php/jkdn>