

Analisis Performa Algoritma Klasifikasi pada Sentimen Ulasan Pengguna terhadap Aplikasi Muamalat DIN

Analysis of Classification Algorithm Performance on User Review Sentiment of the Muamalat DIN Application

Ika Romadoni Yunita¹, Wiga Maulana Baihaqi^{*2}, Aulia Shafira Tri Damayanti³, Luthfi Akhaerunnisa⁴

^{1,2,3,4}Universitas Amikom Purwokerto; Jl. Letjend Pol. Soemarto No.127, Purwokerto, (0281) 623321

^{1,3,4}Program Studi Sistem Informasi, Fakultas Ilmu Komputer Universitas Amikom Purwokerto, Purwokerto

²Program Studi Teknologi Informasi, Fakultas Ilmu Komputer Universitas Amikom Purwokerto, Purwokerto

e-mail: ¹ikarom@amikompurwokerto.ac.id, ^{*2}wiga@amikompurwokerto.ac.id, ³auliashafira2801@gmail.com, ⁴khaerunnisaluthfi@gmail.com

Abstrak

Aplikasi perbankan telah menjadi bagian penting dari kehidupan masyarakat modern. Salah satunya adalah aplikasi Muamalat DIN yang diluncurkan oleh Bank Muamalat Indonesia, yang bertujuan untuk mempermudah nasabah dalam melakukan berbagai transaksi dan kegiatan. Ulasan pengguna terhadap aplikasi ini sangat bervariasi, mulai dari komentar positif hingga komentar negatif. Tujuan dari penelitian ini adalah untuk menganalisis sentimen pengguna terhadap aplikasi Muamalat DIN yang merupakan produk perbankan digital yang ditawarkan oleh Bank Muamalat Indonesia. Dengan menggunakan metodologi yang melibatkan pengumpulan data melalui algoritma scrapping, preprocessing, dan klasifikasi seperti Multi Layer Perceptron (MLP), XGBoost, dan LightGBM, penelitian ini memberikan wawasan tentang efektivitas penyeimbangan data SMOTE dibandingkan dengan undersampling. Temuan penelitian menunjukkan bahwa XGBoost bersama dengan SMOTE memberikan kinerja optimal dalam hal klasifikasi. Hasil penelitian ini menyoroti perlunya teknik penyeimbangan data yang tepat untuk meningkatkan akurasi analisis sentimen dalam aplikasi perbankan syariah. Teknik-teknik ini dapat digunakan sebagai dasar strategi dan keputusan peningkatan layanan pelanggan.

Kata kunci—Aplikasi Perbankan, Muamalat DIN, MLP, XGBoost, LightGBM

Abstract

Banking applications have become an integral part of modern society. One such application is Muamalat DIN, launched by Bank Muamalat Indonesia with the aim of facilitating customers in conducting various transactions and activities. User reviews of this application vary widely, ranging from positive to negative comments. The purpose of this study is to evaluate user attitude on reviews of Bank Muamalat Indonesia's digital banking product, the Muamalat DIN application. This research offers insights into the efficacy of the SMOTE balancing technique compared to undersampling by utilizing a methodology that includes data collection via scrapping techniques, data preprocessing, and the application of Multi Layer Perceptron (MLP), XGBoost, and LightGBM classification algorithms. The results show that SMOTE-paired XGBoost works better for sentiment categorization. The study's conclusion

emphasizes the significance of choosing the right data balancing method to increase sentiment analysis's accuracy in Islamic banking applications, which can be used as a foundation for strategies aimed at enhancing customer service and making decisions.

Keywords— *Banking applications, Muamalat DIN, MLP, XGBoost, LightGBM*

1. PENDAHULUAN

Aplikasi perbankan menjadi bagian penting dari kehidupan masyarakat modern di era teknologi informasi yang terus berkembang pesat. Nasabah dapat melakukan transaksi, melihat informasi keuangan, dan mengelola rekening mereka dengan mudah dengan aplikasi perbankan. Bank Muamalat Indonesia, salah satu lembaga keuangan terkemuka di Indonesia, telah meluncurkan aplikasi perbankan inovatif yang disebut Muamalat DIN (*Digital Islamic Network*).

Aplikasi Muamalat DIN memberikan pelanggan akses cepat dan aman yang dapat diakses kapan saja dan dimana saja oleh seluruh pelanggan baik nasabah ataupun non nasabah. Aplikasi Muamalat DIN memiliki fungsionalitas yang canggih dan antarmuka yang *user-friendly* yang menyajikan berbagai fitur seperti fitur finansial, fitur non finansial, dan fitur menarik lainnya [1].

Dalam penggunaan produk atau aplikasi, ulasan pengguna atau konsumen sangat penting dalam meningkatkan suatu produk karena memiliki tujuan untuk membangun, meningkatkan, dan mengevaluasi kualitas suatu produk [2]. Bank Muamalat Indonesia dapat menggunakan ulasan dan komentar pengguna untuk memahami kebutuhan dan harapan pengguna terhadap aplikasi Muamalat DIN. Dalam hal ini, ulasan pengguna dapat digunakan sebagai analisis sentimen untuk mengetahui bagaimana tanggapan pengguna. Ulasan tersebut dapat diambil dari *Google Play* bagi pengguna Android maupun *Appstore* bagi pengguna IOS.

Analisis sentimen adalah proses untuk memahami, mengekstrak, dan menganalisis data teks secara otomatis untuk mendapatkan informasi tentang sentimen dalam sebuah kalimat, baik positif maupun negatif [3]. Selain karena fakta bahwa saat ini ada banyak data teks yang tersedia, penelitian yang menggunakan analisis sentimen semakin populer karena tuntutan banyak orang tentang pendapat publik tentang topik tertentu. *Dataset* yang digunakan memengaruhi proses analisis sentimen dengan pendekatan yang berbeda diperlukan untuk data yang terdiri dari kumpulan kalimat yang sangat panjang [4].

Dalam beberapa tahun terakhir penelitian tentang analisis sentiment telah banyak dilakukan. Sebagai contoh penelitian [5] melakukan penerapan metode *lexicon* untuk menganalisis sentimen ulasan aplikasi KAI Access dan hasil yang didapatkan adalah sentimen negatif menunjukkan presentase yang paling tinggi mencapai 60,01%. Lalu penelitian [6] melakukan analisis sentimen dari Twitter tentang vaksinasi Covid-19 di Indonesia menggunakan algoritma *Naïve Bayes* dan *XGBoost Classifier*. Hasil yang didapatkan dari penelitian ini menunjukkan bahwa *Naïve Bayes Classifier* mendapatkan performa ROC-AUC yang paling baik sebesar 0.95, sedangkan *XGBoost Classifier* mendapatkan performa ROC-AUC sebesar 0.882. Kemudian pada penelitian lain [7] melakukan analisis sentimen *review* film yang diambil dari *Rotten Tomatoes* menggunakan algoritma *Logistic Regression* dan *Feature Selection Information Gain*. Hasil yang didapatkan dari penelitian ini menunjukkan bahwa algoritma *Logistic Regression* yang menggunakan parameter tanpa menggunakan *Feature Selection Information Gain* mendapatkan hasil *f1-score* sebesar 76.50%.

Berdasarkan uraian yang telah dijelaskan sebelumnya, penelitian ini bertujuan untuk menganalisis ulasan aplikasi Muamalat DIN dengan melihat perbandingan performa dari 3 (tiga) model yang akan digunakan. Metode yang akan digunakan untuk penelitian ini

menggunakan algoritma *Multi Layer Perceptron* (MLP), XGBoost dan LightGBM. Untuk *balancing* menggunakan SMOTE (*Synthetic Minority Oversampling Technique*) untuk penanganan kelas tidak seimbang dan untuk mengetahui performa terbaik peneliti menerapkan metode TF-IDF untuk pembobotan, serta *Information Gain* untuk *feature selection*. Diharapkan analisis sentimen ulasan aplikasi Muamalat DIN ini akan membantu industri perbankan meningkatkan layanan aplikasi perbankan dan memberikan pengalaman pelanggan yang lebih baik, serta memberikan kontribusi pada pemahaman terhadap analisis sentimen di lingkungan aplikasi perbankan Islam dan memberikan panduan bagi pengembangan model sentimen lebih lanjut.

2. METODE PENELITIAN

2.1 Analisis Data

Data yang digunakan untuk penelitian ini berasal dari ulasan pengguna aplikasi Muamalat DIN yang diambil dari *Google Play Store*. Total ada 5971 ulasan yang didapatkan sebagai dataset. Data diambil menggunakan teknik *scrapping* dengan bantuan library *google_play_scaper*. Sentimen yang digunakan pada data ulasan diambil berdasarkan rating ulasan. Dimana rating 1-2 sentimen negatif, rating 3 sentimen netral, dan 4-5 sentimen positif. Jadi pada data ulasan tersebut terdapat tiga kategori sentimen yaitu positif, negatif, dan netral.

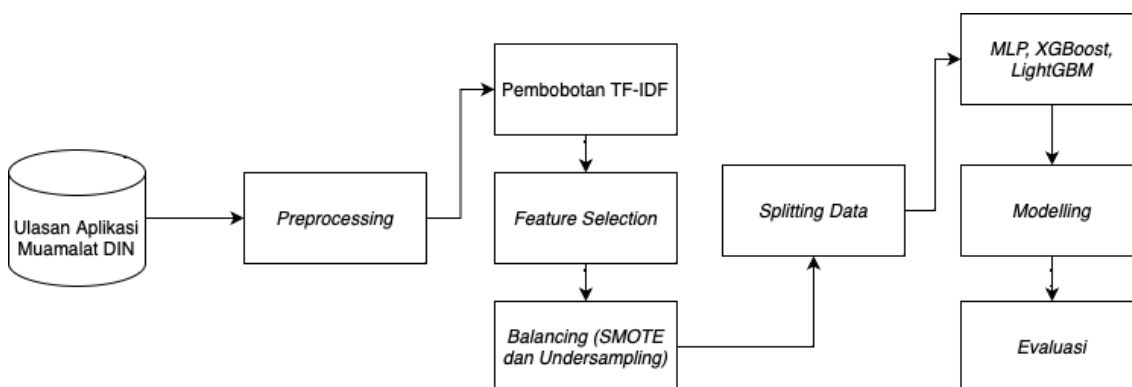
Table 1 Jenis Sentimen

Jenis Sentimen Ulasan	Positif	Netral	Negatif
Muamalat DIN	4667	296	1008

Pada Tabel 1, diketahui data yang mengandung ulasan positif berjumlah 4667 data, dan data yang mengandung ulasan negatif berjumlah 1008, sedangkan data yang mengandung ulasan netral berjumlah 296 data. Data tersebut belum melalui proses *preprocessing* data.

2. 2 Desain Penelitian

Gambar 1 memperlihatkan alur metode penelitian yang digunakan. Dimana proses yang pertama kali dilakukan yaitu mengambil data ulasan aplikasi Muamalat DIN, dilanjutkan dengan *preprocessing* untuk mengelola data yang ada. Kemudian *dataset* dilakukan pembobotan dengan metode TF-IDF, dilanjutkan *balancing* data menggunakan teknik *oversampling* dengan metode SMOTE dan *undersampling* lalu pemilihan fitur dengan metode *Information Gain*. Tahap selanjutnya adalah melakukan pembagian data dengan rasio 80:20, kemudian dilanjutkan pada tahap penerapan model *Multi Layer Perceptron* (MLP), XGBoost dan LightGBM. Proses yang terakhir yaitu akan dibuat *modeling* dan dievaluasi data yang ada.



Gambar 1 Diagram Alir Penelitian

2. 2.1 Preprocessing

Preprocessing adalah teknik penambangan data dengan mengubah data mentah menjadi data yang lebih terstruktur dan mudah dipahami [8]. Langkah-langkah dalam *preprocessing*:

- *Case folding* adalah proses untuk mengubah semua karakter atau teks pada data menjadi huruf kecil, sedangkan karakter yang selain huruf a sampai z akan dihilangkan [9].
- *Tokenizing* merupakan langkah di mana kalimat atau dokumen dibagi menjadi bagian-bagian kecil yang disebut token dan biasanya disertai dengan membuang karakter tertentu seperti tanda baca. [10].
- *Stopwords removal* merupakan proses untuk menghilangkan kata-kata yang tidak memiliki nilai kontribusi banyak pada sebuah dokumen seperti: “ini”, “itu”, “yang”, “ke”, “di”, dan sebagainya [11].
- *Normalization* merupakan proses penghapusan bahasa *alay* dan menormalisasi bahasa singkatan [12]. Setiap kata pada ulasan dinormalisasi dengan mengubah kata yang tidak baku menjadi kata baku [13].
- *Stemming* merupakan proses yang bertujuan untuk mengubah kata-kata yang masih memiliki imbuhan diubah menjadi kata dasar dengan menghilangkan imbuhan di awal dan akhir kata. Proses ini dilakukan secara otomatis menggunakan bantuan *library Sastrawi*. [14].

2. 2.2 Pembobotan (TF-IDF)

Pembobotan kata merupakan proses untuk menentukan bobot pada suatu kata berdasarkan kontribusi dalam sebuah dokumen. TF-IDF (*Term Frequency-Invers Document Frequency*) merupakan metode untuk memberikan bobot pada kata dalam sebuah teks sebanding dengan kemunculan teks tersebut, namun berbanding terbalik dengan teks yang lain. Berikut adalah rumus dari TF-IDF [13].

$$TF = \frac{\text{Word frequency in document}}{\text{Total words in document}} \quad (1)$$

$$IDF = \log \frac{\text{Total number of documents}}{\text{Number of documents containing the word}} \quad (2)$$

2. 2.3 Balancing Data

Metode SMOTE merupakan metode yang digunakan untuk menangani ketidakseimbangan kelas dengan mensintesis kelas baru dari kelas minoritas untuk menyeimbangkan kelas dengan melakukan *resampling* kelas minoritas [8] [15]. Teknik SMOTE bekerja melalui tiga tahap yaitu memilih sampel dari kelas minoritas secara acak, memilih sampel dari k sampel tetangganya, dan membuat sampel baru menggunakan interpolasi linear [16]. Undersampling, sebuah teknik yang digunakan untuk mengurangi ukuran kelas mayoritas, telah banyak digunakan untuk tujuan ini. Ini melibatkan penghapusan instance dari kelas mayoritas, sehingga menciptakan kumpulan data yang lebih seimbang. Studi oleh Na dkk. [17] menyoroti efektivitas kombinasi teknik oversampling (SMOTE) dan undersampling dalam mengatasi ketidakseimbangan data. Pendekatan ini bertujuan untuk menciptakan kumpulan data yang seimbang dengan menangani kelas minoritas dan mayoritas. Lebih lanjut, Rochayani dkk. [18] menekankan prevalensi Random Undersampling (RUS) sebagai metode yang paling banyak digunakan untuk menyeimbangkan data biner. RUS melibatkan pemilihan subset dari instance kelas mayoritas secara acak, sehingga mengurangi jumlahnya agar sesuai dengan

jumlah kelas minoritas. Metode ini terbukti efektif dalam mengatasi kumpulan data yang tidak seimbang, khususnya dalam masalah klasifikasi biner.

2. 2.4 Feature Selection Information Gain

Information Gain merupakan teknik perubahan nilai entropi setelah dilakukan segmentasi pada dataset berdasarkan fitur. Information Gain menghitung seberapa banyak informasi yang relevan berdasarkan fitur kelas pada dataset [13]. Berikut adalah rumus yang digunakan pada seleksi fitur *Information Gain* [7].

$$\begin{aligned}
 IG(t) = & \sum_{i=1}^{|c|} P(c_i) \log P(c_i) \\
 & + P(t) \sum_{i=1}^{|c|} P(c_i|t) \log P(c_i|t) \\
 & + P(\bar{t}) \sum_{i=1}^{|c|} P(c_i|\bar{t}) \log P(c_i|\bar{t})
 \end{aligned} \tag{3}$$

Keterangan:

c_i = kategori

$P(c_i)$ = kategori ganjil

$P(t)$ = kemungkinan bahwa *term*(t) akan muncul dalam dokumen

$P(\bar{t})$ = kemungkinan bahwa *term*(t) tidak muncul dalam dokumen

$P(c_i|t)$ = kondisi kemungkinan kategori akan muncul diberikan *term*(t)

$P(c_i|\bar{t})$ = kondisi kemungkinan bahwa *term*(t) tidak akan muncul

2. 2.5 Pembagian Data

Dalam penelitian ini, pembagian data atau *splitting dataset* dilakukan dengan membaginya menjadi dua bagian yaitu data *train* dan data *test* dengan rasio perbandingan yang digunakan adalah 80:20, artinya 80% untuk data *train* dan 20% untuk data *test*.

2. 2.6 Algoritma Klasifikasi

Pada penelitian ini algoritma yang digunakan diharapkan dapat mewakili beberapa varisai metode klasifikasi. Penelitian ini menggunakan dua algoritma yaitu *Multi Layer Perceptron* (MLP), *LightGBM*, dan *Extreme Gradient Boosting* (XGBoost). MLP adalah salah satu algoritma *deep learning* yang bekerja dengan menggunakan pendekatan *Artificial Neural Network* [16]. *LightGBM* adalah algoritma untuk klasifikasi yang bergantung pada *gradient hoist* [19]. Sedangkan *Extreme Gradient Boosting* (XGBoost) merupakan salah satu algoritma *machine learning* yang dapat diaplikasikan untuk permasalahan klasifikasi dan regresi. XGBoost bekerja dengan menerapkan mekanisme *decision tree* untuk membuat *weak learner* menjadi *learner* yang lebih baik [6] [20].

2. 2.7 Evaluasi

Evaluasi atau pengujian dilakukan pada data *testing* menggunakan model klasifikasi yang telah dibuat. Evaluasi digunakan untuk mengukur keberhasilan klasifikasi, kami menganalisis *confusion matrix* untuk melihat performa *F1-score* dari model klasifikasi apakah presisi dan recall dari model memiliki hasil yang baik. Selanjutnya kami akan menggunakan *cross validation* untuk menghasilkan pengukuran akurasi.

3. HASIL DAN PEMBAHASAN

Pada tahap preprocessing menghasilkan data yang bersih dengan menghilangkan data kosong, dan data duplikat, menghilangkan tanda baca, dan normalisasi untuk menghasilkan data yang bersih, rapi, dan siap untuk dianalisis. Kinerja terbaik akan diterapkan untuk mengklasifikasikan komentar positif, netral, dan negatif pada data ulasan aplikasi Muamalat DIN.

3.1 Dataset

Data yang digunakan pada penelitian ini adalah ulasan pengguna aplikasi Muamalat DIN yang diambil dari *Google Play Store* menggunakan metode *scraping*. Data yang berhasil diambil sebanyak 5971 baris, namun setelah dibersihkan pada tahap *preprocessing* menjadi sebanyak 4194 baris. Kemudian data tersebut akan dibagi dengan rasio 80:20, yakni 80% untuk data *training* dan 20% untuk data *testing*. Tabel 1 menunjukkan contoh data yang digunakan.

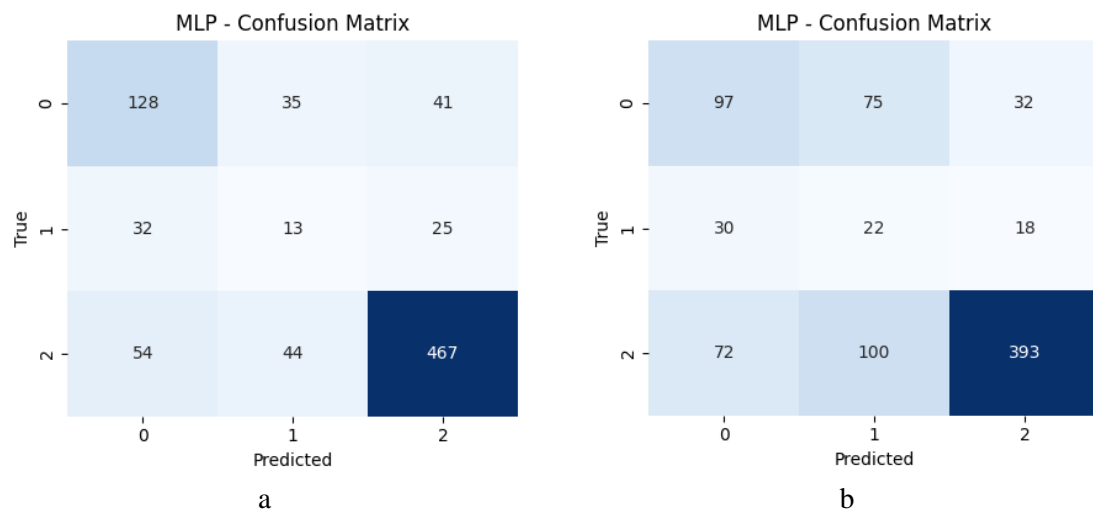
Tabel 1 Sempel *Dataset*

Result	Label
susah mau buka rekening pada tahapan foto ktp tombol memotret memfoto ga berfungsi padahal dah tempat cukup cahaya sesuai frame ayolah aplikasinya bener	Negatif
nyari fitur nya susahhh update jadi ga liat saldo riwayat transaksi udah klik icon ijo aplikasi tetep aja riwayat saldo gak aplikasi nya benar bikinn rumit	Negatif
pernah terjadi eror setelah pengaduan memenuhi kelengkapan data yg dibutuhkan lewat email memang sedikit metepotkan semua pihak terkait aplikasi sangat membantu akhirnya berhasil diatasi saldo yg sempat berkurang eror pembelian pulsa yg gagal eror bukti transaksi tgl yg eror mungkin kembali bertahap kembali rekening terimakasih semua pihak yg membantu semoga makin maju dg inovasi terbaru	Positif
arsitektur interior aplikasinya rumit fitur banyak susah ditemukan ramah pengguna habis transaksi struk transaksi tertampilkan kalo ga langsung disave screen capture bingung nyarinya dimana muamalat mengutamakan haji segala tentang haji aplikasi kayak berceceran dimanamana rapi ditampilkan halaman depan bukan paling penting bukan hal yang benar membantu	Negatif
update terbarunya memang benar kalau mau cek saldo klik icon lampu sebelah kanan ketika buka menu lain info saldo jadi sensor alias jadi bintangbintang jadi tiap buka menu lain selalu pencet icon lampu update sekali pencet icon lampu akan tampil selamanya saldonya update terbaru harus di pencet terus icon lampu otomatis awal	Positif

3.3 Hasil Pengujian Menggunakan MLP, XGBoost dan LightGBM

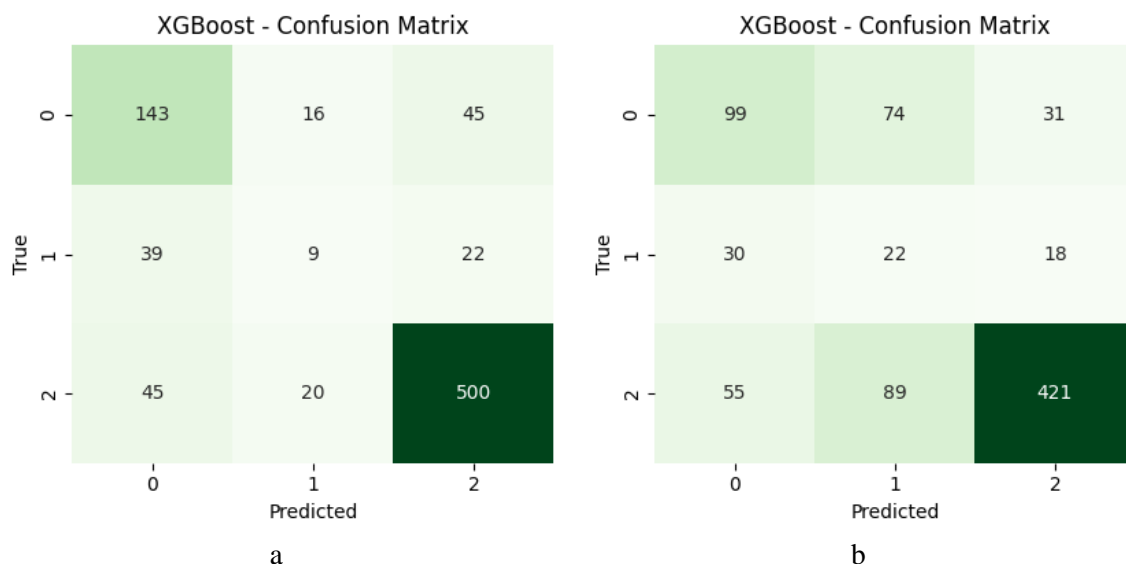
Setelah tahap pra-pemrosesan selesai, tugas berikutnya adalah mengevaluasi hasil yang diperoleh dari algoritma MLP, XGBoost dan LightGBM. Untuk memulai, setiap ulasan diklasifikasikan ke dalam 3 kategori sentiment berdasarkan rating, dengan 1-2 untuk sentimen negatif, 3 untuk sentimen netral dan 4-5 untuk sentiment positif, seperti yang ditunjukkan dalam Tabel 1 yang telah diberi label. Sebelum masuk ke tahap pengujian, dataset akan dilakukan pembobotan dengan metode TF-IDF, lalu diseimbangkan kelasnya dengan metode SMOTE, kemudian dilakukan seleksi fitur dengan metode *Information Gain* dengan nilai $k=500$. Selama tahap pengujian, data dibagi menjadi dua rasio yakni 80:20 untuk pengujian model MLP,

XGBoost dan LightGBM. Setelah itu, data dilatih menggunakan model yang telah ditentukan, yakni MLP, XGBoost dan LightGBM. Selanjutnya, performa model dievaluasi melalui penilaian seperti akurasi, F1-skor/*F-Measure* dan hasil akurasi dari *cross validation*.



Gambar 2 Visualisasi Confusion Matrix MLP (a) SMOTE (b) Undersampling

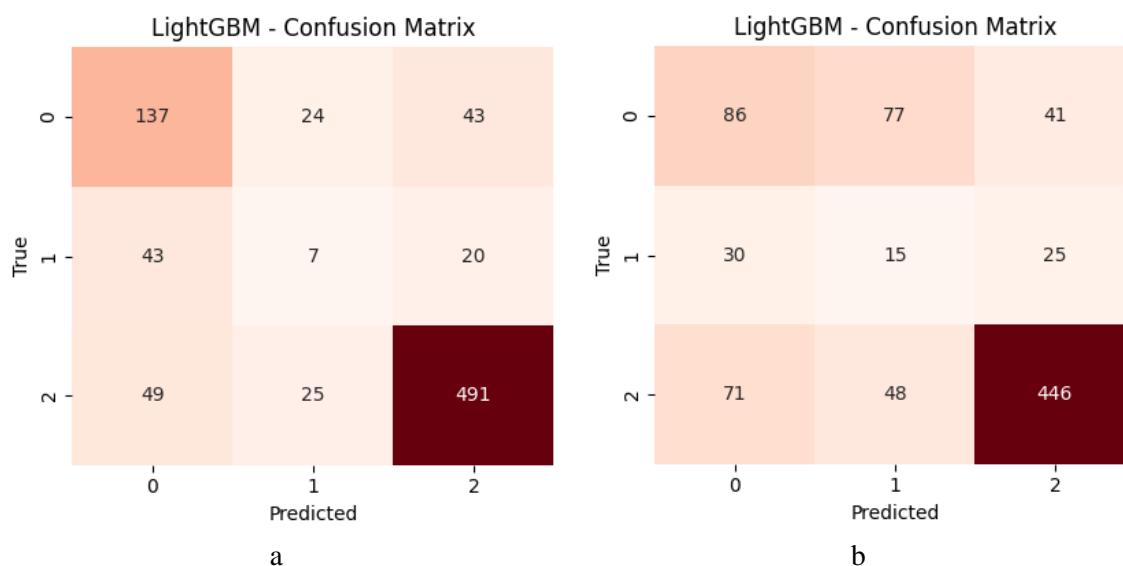
Gambar 2 menunjukkan bahwa model MLP dengan SMOTE mendapatkan hasil prediksi sama dengan hasil actual yaitu pada label kelas 0 (sentimen negatif) sebanyak 128, lalu pada label kelas 1 (sentimen netral) menunjukkan hasil prediksi sama dengan hasil actual sebanyak 13, sedangkan pada label kelas 2 (sentimen positif) menunjukkan hasil prediksi sama dengan hasil actual sebanyak 467. Sedangkan MLP dengan Undersampling mendapatkan hasil prediksi sama dengan hasil actual yaitu pada label kelas 0 (sentimen negatif) sebanyak 97, lalu pada label kelas 1 (sentimen netral) menunjukkan hasil prediksi sama dengan hasil actual sebanyak 22, sedangkan pada label kelas 2 (sentimen positif) menunjukkan hasil prediksi sama dengan hasil actual sebanyak 393.



Gambar 3 Visualisasi Confusion Matrix XGBoost (a) SMOTE (b) Undersampling

Gambar 3 menunjukkan bahwa model XGBoost dengan SMOTE mendapatkan hasil prediksi sama dengan hasil actual yaitu pada label kelas 0 (sentimen negatif) sebanyak 143,

label 1 menunjukkan hasil prediksi sama dengan hasil actual sebanyak 9, sedangkan pada label 2 menunjukkan hasil prediksi sama dengan hasil actual sebanyak 500. Sedangkan XGBoost dengan Undersampling mendapatkan hasil prediksi sama dengan hasil actual yaitu pada label kelas 0 (sentimen negatif) sebanyak 99, label 1 menunjukkan hasil prediksi sama dengan hasil actual sebanyak 22, sedangkan pada label 2 menunjukkan hasil prediksi sama dengan hasil actual sebanyak 421.



Gambar 4 Visualisasi *Confusion Matrix* LightGBM (a) SMOTE (b) Undersampling

Gambar 4 menunjukkan bahwa model LightGBM mendapatkan hasil prediksi sama dengan hasil actual yaitu pada label kelas 0 (sentimen negatif) sebanyak 137, lalu pada label 1 (sentimen netral) menunjukkan hasil prediksi sama dengan hasil actual sebanyak 7, sedangkan label 2 (sentimen positif) menunjukkan hasil prediksi sama dengan hasil actual sebanyak 491. Sedangkan LightGBM dengan Undersampling mendapatkan hasil prediksi sama dengan hasil actual yaitu pada label kelas 0 (sentimen negatif) sebanyak 86, lalu pada label 1 (sentimen netral) menunjukkan hasil prediksi sama dengan hasil actual sebanyak 15, sedangkan label 2 (sentimen positif) menunjukkan hasil prediksi sama dengan hasil actual sebanyak 446. Kemudian pada Tabel 2 menunjukkan hasil evaluasi *f1-score*, akurasi dan *cross validation*.

Tabel 2 Hasil Evaluasi Model

	<i>Balancing Data</i>							
	<i>SMOTE</i>				<i>Undersampling</i>			
Algoritma Klasifikasi	F1-Score	Akurasi	<i>Sensitivity</i>	<i>Spesivisity</i>	F1-Score	Akurasi	<i>Sensitivity</i>	<i>Spesivisity</i>
MLP	0,74	0,72	0,54	0,84	0,47	0,61	0,49	0,80
XGBoost	0,77	0,78	0,57	0,85	0,68	0,64	0,64	0,82
LightGBM	0,75	0,75	0,54	0,85	0,67	0,65	0,65	0,81

Pada Tabel 2 menyajikan hasil analisis sentimen komentar pengguna pada aplikasi Muamalat DIN di Play Store menggunakan tiga algoritma klasifikasi berbeda: MLP (Multilayer Perceptron), XGBoost, dan LightGBM. Setiap algoritma dioptimalkan menggunakan dua teknik penyeimbangan data yang berbeda: SMOTE (Synthetic Minority Over-sampling Technique) dan Undersampling. Menggunakan SMOTE, MLP mencapai Skor F1 0,74 dan akurasi 0,72. Nilai sensitivitas dan spesifisitas yang dihitung masing-masing adalah 0,54 dan 0,84. Namun,

ketika menerapkan undersampling, kinerja MLP ditandai dengan Skor F1 sebesar 0,47 dan akurasi 0,61, serta sensitivitas dan spesifisitas yang berkisar antara 0,49 hingga 0,80. Saat SMOTE digunakan, XGBoost menunjukkan peningkatan kinerja dibandingkan dengan MLP, dengan F1-Score 0,77 dan akurasi 0,78. Sensitivitas dan spesifisitas juga terlihat kurang optimal, dengan nilai rata-rata 0,57 dan 0,85. Undersampling memungkinkan XGBoost mempertahankan kinerja yang relatif tinggi dengan Skor F1 0,68, akurasi 0,64, sensitivitas 0,64, dan spesifisitas 0,82. LightGBM mencapai Skor F1 0,75 dan akurasi 0,75 dengan SMOTE, bersama dengan sensitivitas 0,54 dan spesifisitas 0,85. Dalam kondisi Undersampling, LightGBM melaporkan F1-Score sebesar 0,67 dan akurasi 0,65, sedangkan nilai Sensitivitas dan Spesifisitas masing-masing sebesar 0,65 dan 0,81. Hasil tersebut menunjukkan adanya penurunan performa SMOTE namun hasil yang terbilang kompetitif jika dibandingkan dengan hasil MLP. Secara umum XGBoost dengan SMOTE menghasilkan performa yang optimal dari segi F1-Score dan Akurasi. Teknik penyeimbangan data SMOTE yang konsisten memberikan hasil yang lebih baik daripada undersampling untuk semua algoritma yang dievaluasi, yang menunjukkan efektivitasnya dalam memitigasi distribusi sampel yang tidak merata pada kumpulan data ini.

4. KESIMPULAN

Penelitian ini mengkonfirmasi hipotesis bahwa efektivitas model analisis sentimen dipengaruhi secara signifikan oleh penggunaan teknik penyeimbangan data yang dapat diandalkan. Eksperimen menunjukkan bahwa penggunaan SMOTE secara konsisten meningkatkan kinerja klasifikasi jika dibandingkan dengan undersampling. Hasil penelitian ini mendukung tujuan penelitian untuk mengidentifikasi pendekatan yang paling efektif untuk menganalisis sentimen pengguna pada aplikasi DIN bank syariah Muamalat. Di antara model yang dievaluasi, XGBoost dan SMOTE memberikan hasil terbaik, menekankan efektivitasnya dalam menangkap dan mengklasifikasikan masukan pengguna dengan presisi tinggi.

Semua ini berpotensi sangat membantu Bank Muamalat Indonesia dalam memahami dan merespons secara akurat kebutuhan dan kekhawatiran para penggunanya, yang pada akhirnya akan memperkuat ikatan antar pengguna dan meningkatkan loyalitas mereka. Untuk menyelidiki lebih lanjut hal ini, disarankan untuk menggabungkan teknik penyeimbangan data yang berbeda dan menilai dampak teknik ini pada kumpulan data yang lebih heterogen serta pada lingkungan aplikasi bank yang berbeda untuk memvalidasi dan meningkatkan kegunaan data ini.

5. SARAN

Untuk penelitian lebih lanjut di masa depan, dapat digunakan berbagai metode klasifikasi dan perbandingan terhadap metode lain, juga algoritma yang lain yang dapat memberikan performa yang baik dalam mengatasi data yang tidak seimbang agar dapat meminimalkan kesalahan dalam proses prediksi, serta dapat dilakukan perbandingan dengan *feature selection* yang lainnya.

DAFTAR PUSTAKA

- [1] B. Muamalat, "Bank Muamalat." <https://www.bankmuamalat.co.id/index.php/e-banking/muamalat-din-digital-islamic-network>
- [2] Wijanarto and S. P. Brilianti, "Peningkatan performa analisis sentimen dengan Resampling dan Hyperparameter pada ulasan aplikasi BNI Mobile," *J. Eksplora Inform.*, vol. 9, no. 2, pp. 140–153, 2020, doi: 10.30864/eksplora.v9i2.333.

- [3] G. A. Sandag and J. Waworundeng, “Analisis Sentimen Masyarakat Terhadap Exchange Tokocrypto Pada Twitter Menggunakan Metode LSTM,” *CogITO Smart J.*, vol. 8, no. 2, pp. 411–421, 2022, doi: 10.31154/cogito.v8i2.418.411-421.
- [4] W. Widayat, “Analisis Sentimen Movie Review menggunakan Word2Vec dan metode LSTM Deep Learning,” *J. Media Inform. Budidarma*, vol. 5, no. 3, pp. 1018–1026, 2021, doi: 10.30865/mib.v5i3.3111.
- [5] R. D. Wahyuni and A. N. Utomo, “Penggunaan metode Lexicon untuk analisis sentimen pada ulasan aplikasi KAI Access di Google Play Store,” *J. Rekayasa Inf.*, vol. 11, no. 2, pp. 134–145, 2022.
- [6] A. Irwanto and L. Goeirmanto, “Sentiment Analysis from Twitter about Covid-19 Vaccination in Indonesia using Naïve Bayes and XGboost Classifier Algorithm,” *Sinergi (Indonesia)*, vol. 27, no. 2, pp. 145–152, 2023, doi: 10.22441/sinergi.2023.2.001.
- [7] A. J. Abimanyu, M. Dwifabri, and W. Astuti, “Sentiment Analysis on Movie Review from Rotten Tomatoes Using Logistic Regression and Information Gain Feature Selection,” *Build. Informatics, Technol. Sci.*, vol. 5, no. 1, pp. 162–170, 2023, doi: 10.47065/bits.v5i1.3595.
- [8] S. G. Barus, “Klasifikasi sentimen data tidak seimbang menggunakan algoritma SMOTE dan K-Nearest Neighbor pada ulasan pengguna aplikasi pedulilindungi,” in *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, Jakarta, Indonesia, 2022, pp. 162–173.
- [9] D. F. Zhafira, B. Rahayudi, and Indriati, “ANALISIS SENTIMEN KEBIJAKAN KAMPUS MERDEKA MENGGUNAKAN NAIVE BAYES DAN PEMBOBOTAN TF-IDF BERDASARKAN KOMENTAR PADA YOUTUBE,” *J. Sist. Informasi, Teknol. Informasi, dan Edukasi Sist. Inf.*, vol. 2, no. 1, pp. 55–63, 2021.
- [10] A. F. Sabily, P. P. Adikara, and M. A. Fauzi, “Analisis Sentimen Pemilihan Presiden 2019 pada Twitter menggunakan Metode Maximum Entropy,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 5, pp. 4204–4209, 2019.
- [11] I. N. Fakhri, Jondri, and R. F. Umbara, “Analisis Sentimen pada Kuisioner Kepuasan Terhadap Layanan dan Fasilitas Kampus Universitas Dengan Menggunakan Klasifikasi Support Vector Machine (SVM),” in *e-Proceeding of Engineering*, 2019, pp. 8682–8691.
- [12] N. L. P. C. Savitri, R. A. Rahman, R. Venyutzky, and N. A. Rakhmawati, “Analisis Klasifikasi Sentimen Terhadap Sekolah Daring pada Twitter Menggunakan Supervised Machine Learning,” *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 4, pp. 47–58, 2021.
- [13] M. Falih, N. H. Matondang, and N. Chamidah, “Seleksi fitur information Gain pada analisis sentimen terhadap ulasan aplikasi Flip dengan algoritma Support Vector Machine,” in *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, Jakarta, Indonesia, 2022, pp. 317–326.
- [14] W. Parasati, F. A. Bachtiar, and N. Y. Setiawan, “Analisis Sentimen Berbasis Aspek pada Ulasan Pelanggan Restoran Bakso President Malang dengan Metode Naïve Bayes Classifier,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 4, pp. 1090–1099, 2020.

- [15] C. Cahyaningtyas, Y. Nataliani, I. R. Widiyasari, F. T. Informasi, U. Kristen, and S. Wacana, "Analisis sentimen pada rating aplikasi Shopee menggunakan metode Decision Tree berbasis SMOTE," *AITI J. Teknol. Inf.*, vol. 18, no. 2, pp. 173–184, 2021.
- [16] A. Nurhopipah and C. Magnolia, "Perbandingan metode Resampling pada Imbalanced Dataset untuk klasifikasi komentar program MBKM," *JUPIKOM*, vol. 1, no. 2, 2022.
- [17] Y. Na, L. She, J. Liu, & L. Changjie, "Assessment of equipment operation state with improved random forest", *International Journal of Rotating Machinery*, vol. 2021, p. 1-10, 2021. <https://doi.org/10.1155/2021/8813443>
- [18] M. Rochayani, U. Sa'adah, & A. Astuti, "Finding biomarkers from a high-dimensional imbalanced dataset using the hybrid method of random undersampling and lasso", *Comtech Computer Mathematics and Engineering Applications*, vol. 11, no. 2, p. 75-81, 2020. <https://doi.org/10.21512/comtech.v11i2.6452>
- [19] L. R. Guarneros-Nolasco, N. A. Cruz-Ramos, G. Alor-Hernández, L. Rodríguez-Mazahua, and J. L. Sánchez-Cervantes, "Identifying the main risk factors for cardiovascular diseases prediction using machine learning algorithms," *Mathematics*, vol. 9, no. 20, 2021, doi: 10.3390/math9202537.
- [20] A. Ikegami, I. D. Made, and B. Atmaja, "Analisis sentimen dan pemodelan topik ulasan aplikasi Noice menggunakan XGBoost dan LDA," *JNATIA J. Nas. Teknol. Inf. dan Apl.*, vol. 1, no. November, pp. 325–336, 2022.