Department of Digital Business
Journal of Artificial Intelligence and Digital Business (RIGGS)
Homepage: https://journal.ilmudata.co.id/index.php/RIGGS
Vol. 5 No. 2 (2026) pp: 11786-11795
P-ISSN: 2963-9298, e-ISSN: 2963-914X

Analisis Komparatif Model Machine Learning Berbasis Boosting dalam Analisis Sentimen Ulasan Pengguna Neobank di Google Play Store

Dwi Lufinatul Alifah[1], Virgania Sari[2]
Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Semarang
[1*]dwilufinatulalifah@students.unnes.ac.id

Abstrak

Perkembangan layanan perbankan digital terutama pada aplikasi neobank di Indonesia mendorong meningkatnya jumlah ulasan pengguna di platform Google Play Store, sehingga dibutuhkan metode yang efektif untuk menganalisis opini tersebut secara otomatis. Ulasan pengguna menjadi sumber penting untuk memahami kepuasan dan keluhan nasabah. Penelitian ini bertujuan untuk membandingkan performa algoritma machine learning berbasis boosting, yaitu XGBoost, LightGBM, dan CatBoost, dalam klasifikasi sentimen ulasan pengguna aplikasi Neobank. Data sebanyak 6.963 ulasan dikumpulkan melalui teknik web scraping pada periode Oktober 2025 hingga Februari 2026. Data tersebut kemudian diproses melalui tahapan text preprocessing yang meliputi case folding, pembersihan karakter, tokenisasi, penghapusan stopword, dan stemming menggunakan Sastrawi. Pelabelan sentimen dilakukan secara otomatis dengan pendekatan berbasis leksikon menggunakan kamus InSet (Indonesia Sentiment Lexicon) ke dalam kelas positif dan negatif. Pembobotan kata menggunakan TF-IDF dengan kombinasi unigram dan bigram. Untuk mengatasi ketidakseimbangan kelas, diterapkan teknik SMOTE hanya pada data latih di setiap fold. Evaluasi model dilakukan menggunakan 5-Fold Cross Validation dengan pembagian data 80:20. Hasil pengujian menunjukkan bahwa XGBoost mencapai performa terbaik dengan akurasi rata-rata sebesar 94,30%, presisi 94,35%, recall 94,51%, dan F1-Score 94,37%, diikuti CatBoost dengan akurasi 91,25% dan LightGBM dengan akurasi 91,16%. Seluruh model ini menunjukkan performa di atas 90%, menandakan efektivitas pendekatan yang digunakan. Keunggulan XGBoost diduga karena kemampuannya menangani data sparse dari TF-IDF serta regularisasi yang baik. Ketiga algoritma terbukti mampu mengklasifikasikan sentimen ulasan pengguna Neobank secara akurat, dengan XGBoost sebagai model yang paling unggul dalam pengujian ini.

Kata Kunci: *Analisis Sentimen, Neobank, Xgboost, Lightgbm, Catboost, Tf-Idf, Smote*

1. Latar Belakang

Perkembangan teknologi informasi dan komunikasi telah membawa perubahan besar pada sektor perbankan di Indonesia. Transformasi digital yang didukung oleh meningkatnya penggunaan internet dan *smartphone* mendorong masyarakat untuk memanfaatkan layanan keuangan berbasis digital secara lebih praktis dan efisien. Layanan perbankan yang sebelumnya dilakukan secara konvensional kini beralih menjadi layanan daring yang dapat diakses kapan saja dan di mana saja. Perubahan tersebut turut mempengaruhi perilaku masyarakat dalam melakukan transaksi keuangan, terutama melalui penggunaan *mobile banking* dan internet *banking* [1] . Selain itu, dukungan pemerintah melakukan implementasi National Payment Gateway dan QRIS semakin mempercepat pertembuhan transaksi non-tunai di Indonesia [2].

Seiring dengan perkembangan tersebut, telah muncul berbagai aplikasi layanan perbankan digital seperti Seabank, Bank Neo Commerce, Jenius, dan Bank Jago. Platform – platform ini menawarkan model layanan perbankan tanpa jaringan kantor cabang *(branchless banking)*, dengan keunggulan berupa kemudahan dalam pembukaan rekening secara digital, antarmuka yang intuitif, serta akses terhadap berbagai layanan produk keuangan yang terintegrasi dalam satu aplikasi. Keunggulan platform tersebut menjadi daya tarik utama bagi masyarakat, khususnya segmen generasi muda yang terbiasa dengan era digital [3]. Tingginya persaingan di antara layanan bank digital menjadikan kualitas layanan dan kepuasan pengguna sebagai kunci pengembangan bisnis. Ulasan di Google Play Store menjadi sumber berharga untuk mencerminkan pengalaman pengguna mengenai kualitas layanan, kendala teknis, serta fitur aplikasi [4].

Namun demikian, volume ulasan yang besar dan tidak terstruktur menjadikan analisis yang dilakukan secara manual tidak lagi memadai dalam menghasilkan suatu analisis yang komprehensif dan efisien. Analisis sentimen dilakukan sebagai solusi berbasis *Natural Language Processing* (NLP) dalam melakukan pemrosesan teks ulasan

secara otomatis dan sistematis. Metode ini mengklasifikasikan opini pengguna ke dalam kategori sentimen positif, negatif, atau netral untuk mengukur tingkat kepuasan terhadap layanan bank digital secara kuantitatif [5]. Penerapan analisis sentimen tidak hanya terbatas pada klasifikasi opini, tetapi juga mampu mengidentifikasi aspek layanan yang paling diuatamakan sebagai pengguna, sehingga dapat memberikan suatu landasan berbasis data bagi pengembangan aplikasi perbankan digital. Penelitian ini menerapkan tiga algoritma, yakni XGBoost, LightGBM, dan CatBoost dimana masing – masing algoritma menunjukkan performa baik dalam berbagai studi klasifikasi teks. Perbandingan kinerja ketiga algoritma tersebut dilakukan untuk mengidentifikasi pendekatan yang paling optimal dalam klasifikasi sentimen ulasan pengguna aplikasi Neobank di Indonesia.

2. Kajian Literatur

2.1 Text Mining

Text Mining merupakan suatu pendekatan yang digunakan dalam proses penggalian data untuk memenuhi kebutuhan informasi dengan memanfaatkan teknik *machine learning, data mining, dan natural language processing,* manajemen pengetahuan, serta pencarian informasi.

2.2 Analisis Sentimen

Analisis sentimen *(Opinion Mining)* merupakan suatu bidang dalam pemrosesan bahasa alami *Natural Language Processing* yang bertujuan untuk mengidentifikasi, mengekstraksi, dan mengklasifikasi opini atau emosi yang terkandung dalam teks [6]. Pada analisis sentimen umumnya di bagi menjadi tiga kategori yaitu positif, negatif, dan netral yang menggambarkan kecenderungan opini terhadap suatu objek atau topik tertentu. Pendekatan pada analisis sentimen mengubah opini subjektif menjadi data kuantitatif yang nantinya dapat dianalisis, sehingga dapat memungkinkan pihak peneliti dan praktisi dapat memahami kecenderungan opini publik serta melihat tren perubahan persepsi seiring berjalannya waktu.

2.3 Machine Learning

Machine learning cabang kecerdasan buatan (AI) yang memungkinkan komputer belajar dari data dan meningkatkan kemampuan secara mandiri. Secara sederhana, *machine learning* merupakan proses pengembangan model berbasis data historis untuk menghasilkan prediksi terhadap data baru melalui prinsip optimasi dan statistik [7].

Berdasarkan Pandia [8], *Machine Learning* secara umum dikelompokkan ke dalam empat bagian diantaranya:

a. *Supervised Learning* merupakan metode pembelajaran yang menggunakan data berlabel, dimana input sudah dipasangkan dengan target output yang jelas.
b. *Unsupervised Learning* merupakan metode *machine learning* yang tidak menggunakan data berlabel.
c. Semi-Supervised Learning merupakan perpaduan antara *supervised learning* dan *unsupervised learning.*
d. Reinforcement Learning merupakan metode *machine learning* yang menggunakan sistem *reward* dan *penalti*.

2.4 Text Preprocessing

Text Preprocessing adalah suatu tahap awal yang sangat penting dalam *text mining* yang bertujuan untuk mempersiapkan data teks mentah agar selanjutnya dianalisis oleh algoritma. Selanjutnya, terdapat beberapa tahapan umum dalam text preprocessing diantaranya:

a. *Case Folding* adalah tahap untuk menyeragamkan huruf dengan mengubah huruf kapital menjadi huruf kecil.
b. *Cleansing* merupakan tahap untuk membersihkan karakter tidak diperlukan seperti "@", "#", URL, tanda baca, dan emoticon.
c. Normalisasi adalah proses mengubah kata tak baku menjadi bentuk kata baku.
d. Tokenisasi yaitu proses memecah kalimat menjadi beberapa kata yang disebut token.
e. *Stopword Removal* adalah suatu tahapan preprocessing dengan menghilangkan kata – kata yang kurang penting dalam analisis sentimen.
f. *Stemming* merupakan proses mengubah kata menjadi bentuk dasar dengan menghilangkan imbuhan awalan, sisipan, dan akhiran.

2.5 Pelabelan

DOI: https://doi.org/10.31004/riggs.v5i2.10715

Pelabelan *(labelling)* dalam analisis sentimen merupakan proses pemberian kategori sentimen seperti positif dan negatif pada data teks sebelum tahap klasifikasi. Proses pelabelan menentukan kemampuan model dalam mengenali opini. Pelabelan dapat dilakukan secara manual ataupun secara otomatis menggunakan metode berbasis *lexicon.*

2.6 Stratified K-fold Cross Validation

Stratified K-Fold Validation merupakan teknik salah satu teknik untuk membagi data set menjadi training dan testing. Metode ini digunakan untuk meminimalkan bias dari pembagian sampel acak dengan membagi data ke dalam *k* bagian *(fold),* lalu pelatihan dan pengujian dilakukan bergantian hingga setiap bagian menjadi data uji. Nilai k menunjukkan jumlah lipatan yang digunakan [9].

2.7 TF – IDF

Algoritma *Term Frequecy – Inverse Document Frequency* adalah metode pembobotan kata untuk merepresentasikan dokumen secara numerik dalam *text mining* dan klasifikasi teks, mengukur kepentingan suatu kata dalam dokumen [10].

Term Frequency (TF) didefinisikan sebagai frekuensi relatif term dalam suatu dokumen, yang dinyatakan sebagai:

$$TF_{ij} = \frac{f_{ij}}{\sum_{k=1}^{m} f_{kj}} \qquad (1)$$

Inverse Document Frequency (IDF) mengukur tingkat keterbatasan suatu *term* dalam keseluruhan kumpulan dokumen yang dinyatakan sebagai:

$$IDF_i = \log\left(\frac{N}{DF_i}\right) \qquad (2)$$

Bobot akhir TF-IDF diperoleh melalui perkalian kedua komponen tersebut:

$$TF - IDF_{ij} = TF_{ij} \times IDF_i \qquad (3)$$

Melalui pembobotan ini, kata yang sering muncul dalam satu dokumen tetapi jarang muncul pada dokumen lain akan memiliki bobot tinggi dan dianggap kebih representatif dalam klasifikasi teks.

2.8 SMOTE

Synthetic Minority Oversampling (SMOTE) adalah metode *oversampling* yang digunakan untuk menangani ketidakseimbangan kelas pada proses klasifikasi. Ketidakseimbangan kelas terjadi ketika jumlah data pada kelas mayoritas lebih banyak dibanding kelas minioritas sehingga dapat memperngaruhi kinerja model klasifikasi [11].

Secara matematis, jarak antar sampel pada metode SMOTE dihitung menggunakan jarak *Euclidean* yang dirumuskan sebagai berikut:

$$(x_i, x_j) = \sqrt{\sum_{m=1}^{n} (x_{im} - x_{jm})^2} \qquad (4)$$

Perhitungan dilakukan dengan mengukur jarak antara data berdasarkan seluruh fitur untuk menentukan tetangga terdekat. Setelah itu, data sintetis baru dibentuk menggunakan interpolasi linear antara data minoritas dan tentangganya dengan persamaan:

$$x_{new} = x_i + \delta(x_{zi} - x_i) \qquad (5)$$

Tahap ini bertujuan untuk menambah jumlah data pada kelas minoritas sehingga menjadi lebih seimbang.

DOI: https://doi.org/10.31004/riggs.v5i2.10715

2.9 XGBOOST

Extreme Gradient Boosting (XGBOOST) merupakan algoritma *machine learning* berbasis pohon keputusan yang dikembangkan untuk menangani permasalahan klasifikasi dan regresi. Algoritma ini menerapkan model boosting, yaitu membangun model secara bertahap dengan menambahkan pohon keputusan secara berurutan. Setiap pohon baru bertugas memperbaiki kesalahan prediksi dari pohon sebelumnya [12]. Persamaan XGBoost dapat ditunjukkan sebegai berikut:

$$L^{(t)} = \sum_{i=1}^{n} l\left(\left(y_i, \hat{y}_i^{(t-1)}\right) + f_t(x_i)\right) + \Omega(f_t) \quad (6)$$

Pada persamaan tersebut, n menunjukkan jumlah model yang digunakan, l merupakan fungsi *loss* untuk mengukur selisih anatara nilai aktual dan prediksi, $f_t(x_i)$ adalah model baru yang dibangun pada iterasi ke-t , sedangkan $\Omega(f_t)$ merupakan fungsi regularisasi untuk mencegah *overfitting.*

2.10 LIGHTGBM

Light Gradient Boosting Machine adalah model *machine learning* berbasis pohon keputusan yang menawarkan efisiensi tinggi dalam pemrosesan data berskala masif, dengan keunggukan kecepatan yang optimal serta penggunaan memori yang rendah. Persamaan pada model LightGBM pada iterasi ke-m dirumuskan sebagai:

$$F_m(x) = F_{m-1}(x) + \eta h_m(x) \qquad (7)$$

Persamaan diatas menunjukkan bahwa model prediksi ke-m yaitu $F_m(x)$ serta diperoleh dari prediksi sebelumnya $F_{m-1}(x)$ yang nantinya diperbaiki menggunakan pohon keputusan baru $h_m(x)$ dengan pengaruh oleh η. Setiap pohon baru berfungsi untuk mengoreksi kesalahan iterasi sebelumnya sehingga performa model meningkat secara bertahap.

2.11 CATBOOST

Categorikal Boosting adalah algoritma *machine learning* yang termasuk dalam kategori *gradient boosting,* di mana pohon keputusan biner berfungsi sebagai *base laner* dasarnya. Algoritma ini memiliki kemampuan dalam memproses fitur kategorikal serta menekan risiko overfitting dengan mengaplikasikan estimasi *ordered boosting* [13]. Secara sistematis, fungsi objektif pada Catboost dapat dirumuskan sebagai berikut:

$$obj(\theta) = \sum_{i=1}^{n} L\,(y_i, \hat{y}_i) + \lambda \sum_{j=1}^{T} \Omega(\,f_j)$$

Pada persamaan tersebut, komponen fungsi kerugian $L\,(y_i, \hat{y}_i)$ bertujuan untuk meminimalkan kesalahan prediksi, sedangkan fungsi regulasi $\Omega(f_j)$ menekan kompleksitas model agar tidak terjadi *overfitting.*

2.12 Confusion Matrix

Confusion matrix merupakan metode evaluasi yang digunakan untuk mengukur kinerja model klasifikasi dengan membandingkan hasil prediksi model terhadap kelas sebenarnya. Confusion matrix dapat dibagi menjadi empat komponen, diantaranya:

a. TP (*True Positive*) : jumlah data yang diprediksi benar sebagai positif.
b. TN (*True Negative*) : jumlah data negatif yang diprediksi benar sebagai negatif.
c. FP (False Positive) : jumlah data negatif tetapi diprediksi positif.
d. FN (False Negative) : jumlah data positif tetapi diprediksi negatif.

Struktur tabel confusion matrix dapat digambarkan:

Tabel 1 *Confusion Matrix*

DOI: https://doi.org/10.31004/riggs.v5i2.10715

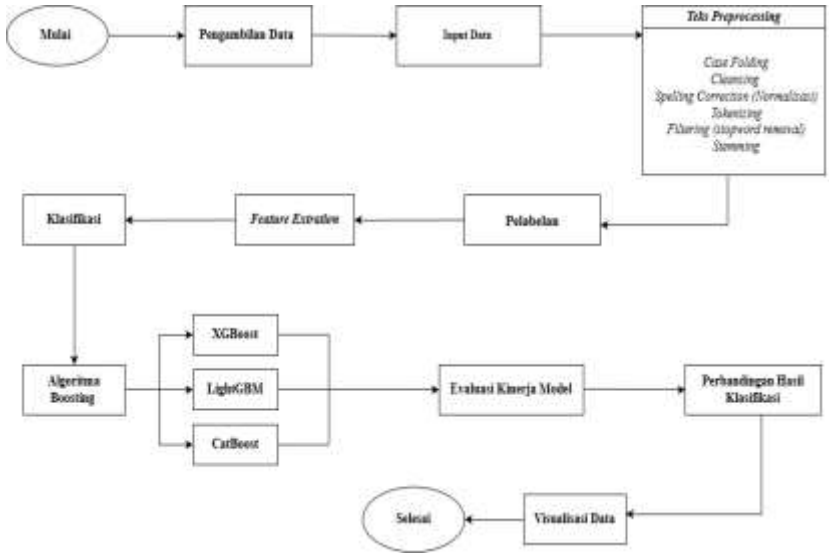
	Prediksi Positif	Prediksi Negatif
Aktual Positif	TP	FN
Aktual Negatif	FP	TN

2.13 Word Cloud

Word Cloud merupakan suatu teknik visualisasi yang menyajikan representasi visual tentang frekuensi kemunculan kata dalam suatu teks, dengan kata – kata yang sering muncul pada tampilan dalam ukuran yang lebih besar [14]. Dalam analisis sentimen, *Word Cloud* berfungsi sebagai alat untuk visualisasi yang memungkinkan peneliti untuk melihat dengan cepat di mana kata – kata yang mencerminkan sentimen positif dan negatif terhadap suatu produk atau layanan.

3. Metode Penelitian

Pendekatan kuantitatif digunakan untuk menganalisis dan memandingkan performa algoritma XGBoost [15], LightGBM, dan CatBoost dalam klasifikasi sentimen ulasan pengguna aplikasi Neobank. Data yang digunakan merupakan data sekunder berupa ulasan pengguna aplikasi Neobank di Google Play Store yang dikumpulkan melalui teknik *web scraping* menggunakan *library google-play-scraper*. Data diambil pada periode 6 Oktober 2025 – 16 Februari 2026 dengan total 6.963 ulasan yang mencakup rating 1 – 5. Selanjutnya, data diproses melalui tahap *preprocessing* dan digunakan untuk pelatihan serta evaluasi model menggunakan nilai akurasi, presisi, recall, dan F1-Score untuk menentukan algoritma dengan performa terbaik. Tahapan dalam penelitian ini ditampilkan melalui gambar 1 berikut.



Gambar 1 Distibusi Sentimen

4. Hasil dan Pembahasan

4.1 Scraping Data

Data penelitian ini diperoleh melalui teknik *web scraping* dari ulasan pengguna aplikasi bank digital Neobank di Google Play Store. Proses scraping dilakukan menggunakan *Google Colaboratory* dengan bahasa pemograman python serta parameter lang='id'dan country='id'. Hasil dari proses *scraping* memperoleh 6.963 ulasan dan data ulasan tersebut dikonversi ke dalam format DataFrame dan disimpan dalam file CSV. Teknik scraping tersebut juga menghasilkan ulasan meliputi review_id, user_name, at, score, content, likes, app_version, replyContent, dan repliedAt.

4.2 Text Preprocessing Data

Data ulasan hasil dari proses web scraping masih berupa data mentah sehingga perlu dilakukan preprocessing agar lebih terstruktur dan siap dianalisis. Pada tahap ini, kolom yang tidak relevan dihapus, sedangkan data

DOI: https://doi.org/10.31004/riggs.v5i2.10715

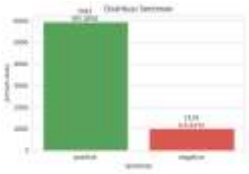
difokuskan pada kolom at, score, dan content. Selanjutnya, dilakukan rename kolom menjadi Date, rating, dan ulasan.Tahapan *Text Preprocessing* diantaranya seperti tabel dibawah ini:

Tabel 3 *Text Preprocessing*

Ulasan	*Case Folding*	*Cleaning*	*Normalization*	*Tokenizing*	*Stopword Removal*	*Stemming*
pengajuan pinjamannya Kenapa susah di acc & di setujui ya...??	pengajuan pinjamannya kenapa susah di acc & di setujui ya...??	pengajuan pinjamannya kenapa susah di acc di setujui ya	pengajuan pinjamannya kenapa susah di setuju di setujui ya	['pengajuan', 'pinjamannya', 'kenapa', 'susah', 'di', 'setuju', 'di', 'setujui', 'ya']	['pengajuan', 'pinjamannya', 'susah', 'setuju', 'setujui']	ngajuan pinjam susah setuju setuju

4.3 Pelabelan

Setelah tahap preprocessing, dilakukan pelabelan sentimen untuk menentukan suatu ulasan termasuk positif atau negatif. Pelabelan dilakukan menggunakan pendekatan leksikon. Proses ini menghitung jumlah positif dan negatif pada hasil ulasan *stemming.*



Gambar 2 Distibusi Sentimen

Pada distribusi sentimen di atas menggambarkan data yang tidak seimbang antara ulasan negatif dan positif. Distribusi data pada kelas positif berjumlah 5945 data sedangkan pada kelas negatif berjumlah 1018 data.

4.4 TF-IDF (Pembobotan Kata)

Data ulasan yang telah melalui tahap pelabelan sentimen masih berbentuk teks sehingga perlu dikonversi menjadi data numerik agar selanjutnya dapat diproses oleh algoritma. Koversi tersebut dilakukan menggunakan TF -IDF yang memberikan bobot pada setiap kata berdasarkan tingkat kemunculannya. Berikut pembobotan TF-IDF dengan nilai tertinggi.

Tabel 4 Pembotan TF -IDF

Kata	Rata – rata TF IDF
bagus	0.148246
mantap	0.089044
sangat	0.076376
aplikasi	0.060201
sekali	0.041723
baik	0.041707

4.5 Split Data dan K-Fold Cross Validation

Dataset dibagi menjadi data latih dan data uji dengan proporsi 80:20, dimana data latih digunakan untuk membangun model sedangkan data uji untuk mengevaluasi kinerja model. Memastikan hasil dari evaluasi yang lebih baik dan tidak bergantung pada suatu pembagian data tertentu, pengujian model dilakukan melalui tahap 5-*Fold Cross Validation*. Metode ini bekerja dengan cara membagi dataset ke dalam lima *fold*, kemudian menjalankan proses pelatihan serta pengujian secara bergantian pada setiap *fold*, sehingga setiap bagian data dapat menjadi data uji tepat satu kali. Adapun hasil yang diperoleh dari proses tesebut disajikan sebagai berikut:

Tabel 4 Hasil Pembagian Split Data

Data	*Jumlah data pada fold ke-*

DOI: https://doi.org/10.31004/riggs.v5i2.10715

	1	2	3	4	5
Data Training	5532	5533	5533	5533	5533
Data Testing	1384	1383	1383	1383	1383

4.6 SMOTE (Synthetic Minority Oversampling)

Teknik SMOTE diterapkan untuk mengatasi masalah ketidaksimbangan kelas dengan menghasilkan data sintetis pada kelas minioritas sehingga distribusi data menjadi seimbang dan model tidak bias terhadap kelas mayoritas. Penerapan SMOTE hanya dilakukan pada data latih disetiap *fold* agar informasi pada data uji tetap memperlihatkan kondisi data asli. Hasil pembagian data setelah penerapan SMOTE disajikan sebagai berikut:

Tabel 5 *Synthetic Minority Oversampling*

Data	*Jumlah data pada fold ke-*				
	1	2	3	4	5
Data Training	9440	9440	9440	9440	9440
Data Testing	1384	1383	1383	1383	1383

4.7 Penerapan Algoritma XGBoost

Klasifikasi sentimen ulasan aplikasi Neobank pada penerapan algoritma *Extreme Gradient Boosting*. XGBoost merupakan algoritma *ensemble learning* berbasis pohon keputusan, dimana setiap pohon baru yang dibentuk berfokus untuk memperbaiki kesalahan prediksi pohon sebelumnya sehingga model menjadi lebih optimal. Hasil klasifikasi menggunakan metode XGBoost sebagai berikut:

Tabel 6 Hasil 5-Fold Cross Validation XGBoost

Fold	*Extreme Gradient Boosting*			
	Accuracy	*Precision*	*Recall*	*F1-Score*
1	0.9449	0.9436	0.9449	0.9441
2	0.9530	0.9518	0.9530	0.9521
3	0.9313	0.9293	0.9313	0.9301
4	0.9530	0.9519	0.9530	0.9510
5	0.9430	0.9410	0.9430	0.9412
Mean	0.9430	0.9435	0.9451	0.9437

Tabel 6 Hasil Penerapan XGBoost

Kelas	*Precision*	*Recall*	*F1-Score*	*Support*
Negative	0.86	0.79	0.83	203
Positive	0.96	0.98	0.97	1180
Accuracy			0.95	1383
Weighted Average	0.95	0.95	0.95	1383

4.8 Penerapan Algoritma LightGBM

Klasifikasi sentimen pada penerapan metode LightGBM yang merupakan algoritma ensemble berbasis pohon keputusan yang menerapkan pendekatan *gradient boosting* dengan strategi pertumbuhan pohon secara *leaf-wise*, berbeda dengan algoritma *boosting* konvensional yang menggunakan strategi level-wise, sehingga proses pelatihan menjadi lebih efisien. Hasil dari klasifikasi menggunakan metode LightGBM sebagai berikut:

Fold	*Light Gradient Boosting Machine*			
	Accuracy	*Precision*	*Recall*	*F1-Score*
1	0.9070	0.9361	0.9070	0.9145
2	0.9151	0.9388	0.9151	0.9213
3	0.9088	0.9368	0.9088	0.9160
4	0.9105	0.9410	0.9105	0.9180
5	0.9168	0.9381	0.9168	0.9226
Mean	0.9116	0.9382	0.9116	0.9185

Tabel 7 Hasil Penerapan LightGBM

DOI: https://doi.org/10.31004/riggs.v5i2.10715

Kelas	*Precision*	*Recall*	*F1-Score*	*Support*
Negative	0.64	0.98	0.77	203
Positive	1.00	0.91	0.95	1180
Accuracy			0.92	1383
Weighted Average	0.94	0.92	0.92	1383

4.9 Penerapan Algoritma CatBoost

Algoritma Catboost berbasis *gradient boosting* dengan teknik *Ordered Boosting* yang memproses data secara berurutan guna meminimalkan *prediction shift*. Catboost membangun pohon simetris dimana setiap level menggunakan kondisi pemisahan yang sama pada seluruh *node*. Hasil klasifikasi menggunakan CatBoost dapat disajikan sebagai berikut:

Tabel 8 Hasil Penerapan CatBoost

Fold	*Categorikal Boosting*			
	Accuracy	*Precision*	*Recall*	*F1-Score*
1	0.8645	0.9273	0.8645	0.8798
2	0.9503	0.9582	0.9503	0.9524
3	0.9024	0.9369	0.9024	0.9110
4	0.9412	0.9554	0.9412	0.9447
5	0.9042	0.9350	0.9042	0.9122
Mean	0.9125	0.9426	0.9125	0.9200

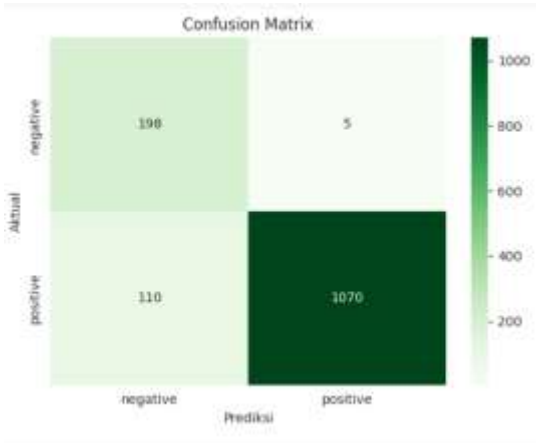
Tabel 8 Hasil Penerapan LightGBM

Kelas	*Precision*	*Recall*	*F1-Score*	*Support*
Negative	*0.65*	*0.97*	*0.78*	*203*
Positive	*0.99*	*0.91*	*0.95*	*1180*
Accuracy		*0.92*		*1383*
Weighted Average	*0.94*	*0.92*	*0.93*	*1383*

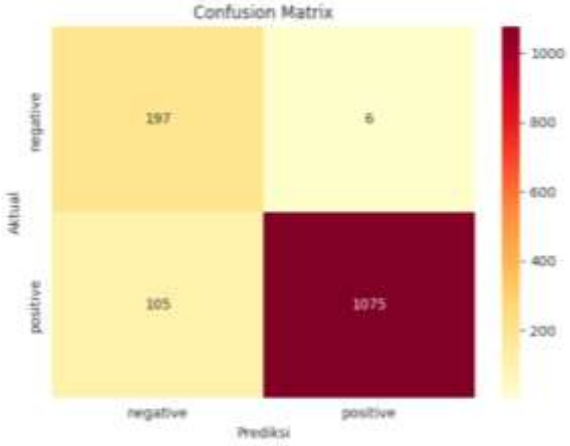
4.10 Evaluasi Model dan Word Cloud

Gambar 3 *Confusion Matrix XGBoost*

DOI: https://doi.org/10.31004/riggs.v5i2.10715



Gambar 4 *Confusion Matrix LightGBM*



Gambar 5 *Confusion Matrix CatBoost*

Gambar 6 *Word Cloud Positi*ve

Gambar 7 *WordCloud Negatifve*

DOI: https://doi.org/10.31004/riggs.v5i2.10715

5. Kesimpulan

Penelitian ini menganalisis sentimen ulasan pengguna aplikasi Neobank di Google Play Store menggunakan tiga algoritma *machine learning* berbasis *boosting*, yaitu XGBoost, LightGBM, dan CatBoost, terhadap 6.963 ulasan yang dikumpulkan melalui *web scraping* dan diproses menggunakan *text preprocessing*, pelabelan leksikon, pembobotan TF-IDF, SMOTE, serta evaluasi 5-*Fold Cross Validation*. Hasil penelitian menunjukkan bahwa XGBoost mencapai performa terbaik dengan akurasi rata-rata 94,30%, unggul dibandingkan CatBoost 91,25% dan LightGBM 91,16%, sehingga ketiga algoritma terbukti mampu mengklasifikasikan sentimen pengguna Neobank dengan tingkat akurasi yang baik. Untuk pengembangan lebih lanjut, disarankan agar penelitian berikutnya mengeksplorasi pendekatan *deep learning* seperti BERT atau IndoBERT yang lebih mampu memahami konteks sistem teks berbahasa Indonesia, memperluas cakupan data dari platform lain, serta menerapkan *hyperparameter tuning* dan anotasi label secara manual untuk meningkatkan kualitas dan performa model secara keseluruhan.

Refrensi

[1] Bank Indonesia, "Laporan Kelembagaan Bank Indonesia Triwulan IV-2023," Jakarta, 2023.
[2] Bank Indonesia, "Tinjauan Kebijakan Moneter September 2023," Jakarta, 2023.
[3] DigitalBank.id, "Survei Ipsos : Bank digital kian jadi andalan, SeaBank,Jago, dan Neo pimpin pasar."
[4] R. Purnamasari and A. R. Thaha, "Understanding user satisfaction in digital banking : Sentiment analysis and topic modeling of Indonesian bank reviews," vol. 19, no. 3, pp. 101–112, 2025, doi: 10.17323/2587-814X.2025.3.101.112.
[5] L. Zhang, S. Wang, and B. Liu, "Deep learning for sentiment analysis: A survey," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 8, no. 4, pp. 0–3, 2025, doi: 10.1002/widm.1253.
[6] T. Wahyudi, N. Purwati, N. Hasan, and G. B. Sulistyo, "Tren NLP dalam Analisis Sentimen Media Sosial : Tinjauan Sistematis dan Bibliometrik," vol. 25, pp. 206–212, 2025.
[7] R. S. Nurhalizah and R. Ardianto, "Analisis Supervised dan Unsupervised Learning pada Machine Learning : Systematic Literature Review," vol. 4, no. 1, pp. 61–72, 2024, doi: https://doi.org/10.54082/jiki.168.
[8] M. Pandia, P. Sihombing, P. Simamora, and R. Kaban, "Kajian Literatur Multimedia Retrieval : Machine Learning Untuk Pengenalan Wajah," vol. 7, pp. 161–166, 2024.
[9] T. Ridwansyah, "Implementasi Text Mining Terhadap Analisis Sentimen Masyarakat Dunia Di Twitter Terhadap Kota Medan Menggunakan K-Fold Cross Validation Dan Naïve Bayes Classifier," vol. 2, no. 5, pp. 178–185, 2022, doi: https://doi.org/10.30865/klik.v2i5.362.
[10] M. Hafizh Mahendra, D. Triantoro Murdiansyah, and K. Muslim Lhaksmana, "Analisis Sentimen Tweet COVID-19 menggunakan K-Nearest Neighbors dengan TF-IDF dan Ekstraksi Fitur CountVectorizer," *Dike J. Ilmu Multidisiplin*, vol. 1, no. 2, pp. 37–43, 2023.
[11] C. Candra, K. W. Chandra, and H. Irsyad, "Efektifitas SMOTE dalam Mengatasi Imbalanced Class Algoritma K-Nearest Neighbors pada Analisis Sentimen terhadap Starlink," *J. Ilmu Komput. dan Inform.*, vol. 4, no. 1, pp. 31–42, 2024, doi: 10.54082/jiki.132.
[12] N. K. P. Indrayuni, N. M. M. R. Desmayani, I. D. A. A. T. Pramawati, I. M. S. Sandhiyasa, and K. K. Widiartha, "Sentiment Analysis on Rupiah Depreciation Against USD Using XGBoost," *J. Appl. Informatics Comput.*, vol. 9, no. 5, pp. 2521–2532, 2025, doi: 10.30871/jaic.v9i5.10751.
[13] N. K. Dewi, *Deteksi Fake Follower Instagram menggunakan Catboost Classifer*. 2021.
[14] V. Ibrahim, "A Word Cloud Model based on Hate Speech in an Online Social Media Environment A Word Cloud Model based on Hate Speech in an Online Social Media Environment," vol. 18, no. 2, 2021.
[15] Sugiyono, *Metode penelitian kuantitatif, kualitatif, dan kombinasi (Mixed methods)*, Edisi Revi. Bandung, Indonesia: Alfabeta, 2022.

DOI: https://doi.org/10.31004/riggs.v5i2.10715