

IMPLEMENTASI ALGORITMA NAÏVE BAYES UNTUK ANALISIS SENTIMEN PEMBELIAN ROBUX PADA MEDIA SOSIAL X

Hanum Surya Herfiana¹, Nina Sulistyowati², Billy Ibrahim Hasbi³

^{1,2,3} Sistem Informasi, Fakultas Ilmu Komputer, Universitas Singaperbangsa Karawang

Jl. HS.Ronggo Waluyo, Puseurjaya, Telukjambe Timur, Karawang

hanum.suryaherfiana@gmail.com¹, nina.sulistio@unsika.ac.id², billy.ibrahim@staff.unsika.ac.id³

Abstrak

Perkembangan industri game online telah mendorong terbentuknya ekosistem ekonomi virtual yang memungkinkan pengguna melakukan berbagai transaksi digital. Salah satu bentuk transaksi tersebut adalah pembelian mata uang virtual Robux pada platform game Roblox. Pembelian Robux dapat dilakukan melalui platform resmi maupun melalui pihak ketiga yang sering menawarkan harga lebih murah. Namun, transaksi melalui platform non-resmi sering menimbulkan berbagai permasalahan seperti risiko penipuan, keamanan transaksi, serta potensi pelanggaran kebijakan platform. Fenomena ini memunculkan berbagai opini dan pengalaman pengguna yang banyak dibagikan melalui media sosial, khususnya pada platform X. Penelitian ini bertujuan untuk menganalisis sentimen pengguna media sosial X terhadap pengalaman pembelian Robux menggunakan pendekatan text mining. Metode yang digunakan dalam penelitian ini adalah kerangka kerja SEMMA (Sample, Explore, Modify, Model, Assess) dengan algoritma Naïve Bayes Classifier serta pembobotan kata menggunakan metode TF-IDF. Dataset penelitian terdiri dari 393 cuitan berbahasa Indonesia yang dikumpulkan pada rentang waktu tahun 2023 hingga 2025. Hasil penelitian menunjukkan bahwa sentimen negatif mendominasi dengan jumlah 190 data, diikuti sentimen netral sebanyak 108 data, dan sentimen positif sebanyak 65 data. Hasil evaluasi menunjukkan bahwa algoritma Naïve Bayes mampu mengklasifikasikan sentimen pengguna dengan cukup baik berdasarkan nilai akurasi, precision, dan recall yang diperoleh dari confusion matrix.

Kata Kunci : Analisis Sentimen, Naïve Bayes, Text Mining, Robux, Roblox

Abstract

The development of the virtual economy in the online gaming industry has increased digital transaction activities, including the purchase of Robux virtual currency on the Roblox platform. Robux purchases can be made through official platforms or third-party sellers that often offer cheaper prices. This phenomenon generates various opinions and experiences shared by users on social media, particularly on platform X. This study aims to analyze user sentiment on social media X regarding Robux purchasing experiences and evaluate the performance of the Naïve Bayes algorithm in classifying these sentiments. The research method used is a text mining approach with the SEMMA framework (Sample, Explore, Modify, Model, Assess). The dataset consists of 393 Indonesian-language tweets collected from social media X between 2023 and 2025. The preprocessing stage includes cleaning, tokenizing, stopword removal, and stemming. Word weighting is then performed using the TF-IDF method. The classification process is conducted using the Naïve Bayes Classifier algorithm and evaluated using a confusion matrix to obtain accuracy, precision, and recall values. The results show that negative sentiment dominates with 190 data, followed by neutral sentiment with 108 data and positive sentiment with 65 data. Negative sentiment is mostly related to security concerns in transactions through unofficial platforms, while positive sentiment is associated with transaction convenience and promotional offers provided by official platforms. These findings indicate that the Naïve Bayes method can be effectively used for sentiment classification in social media text data.

Keyword : Sentiment Analysis, Naïve Bayes, Text Mining, Robux, Social Media

PENDAHULUAN

Perkembangan teknologi informasi telah mendorong terbentuknya berbagai bentuk ekonomi digital, termasuk ekonomi virtual yang berkembang dalam industri game online. Ekonomi virtual merupakan sistem ekonomi yang berlangsung dalam lingkungan digital di mana pengguna dapat melakukan transaksi, menciptakan nilai, serta memperoleh keuntungan melalui aset digital seperti mata uang virtual dan item dalam permainan [1]. Dalam ekosistem permainan daring, ekonomi virtual menjadi bagian penting dari pengalaman pengguna karena memungkinkan pemain untuk

membeli, menjual, atau menukar berbagai item digital yang tersedia dalam permainan [2].

Salah satu platform permainan daring yang cukup populer adalah Roblox, yaitu platform yang memungkinkan pengguna untuk membuat, memainkan, serta membagikan berbagai permainan yang dibuat oleh komunitas pengguna. Dalam ekosistem Roblox terdapat mata uang virtual yang disebut Robux yang digunakan untuk membeli berbagai item digital seperti pakaian avatar, akses permainan, maupun fitur premium lainnya. Keberadaan mata uang virtual ini menunjukkan bahwa aktivitas dalam game tidak hanya terbatas pada hiburan semata, tetapi juga telah berkembang menjadi bagian dari sistem ekonomi digital yang melibatkan transaksi antar pengguna [3].

Dalam praktiknya, pembelian Robux dapat dilakukan melalui berbagai saluran, baik melalui *platform* resmi yang disediakan oleh Roblox maupun melalui platform pihak ketiga yang sering menawarkan harga yang lebih murah. Perbedaan saluran pembelian ini sering menimbulkan berbagai pandangan di kalangan pemain karena berkaitan dengan faktor harga, keamanan transaksi, serta tingkat kepercayaan terhadap penyedia layanan [4]. *Platform* resmi umumnya dianggap lebih aman dan terpercaya, namun memiliki harga yang relatif lebih tinggi dibandingkan dengan platform non-resmi yang sering menawarkan harga lebih murah tetapi memiliki risiko tertentu seperti penipuan atau pelanggaran kebijakan platform.

Di sisi lain, media sosial menjadi salah satu sarana utama bagi pengguna untuk menyampaikan opini dan pengalaman mereka terhadap berbagai layanan digital. Platform X (sebelumnya Twitter) memungkinkan pengguna untuk berbagi komentar, pengalaman, serta diskusi secara real-time mengenai berbagai topik yang sedang berkembang. Interaksi pengguna yang tinggi pada platform ini menjadikannya sebagai salah satu sumber data yang relevan dalam penelitian yang berkaitan dengan perilaku pengguna dan opini publik [5]. Komentar yang dituliskan oleh pengguna sering mencerminkan pengalaman serta persepsi mereka terhadap suatu produk atau layanan digital sehingga dapat dimanfaatkan sebagai sumber data untuk menganalisis sentimen masyarakat [6].

Untuk memahami kecenderungan opini yang muncul dalam data teks media sosial, dapat digunakan metode analisis sentimen. Analisis sentimen merupakan teknik yang digunakan untuk mengekstrak, mengidentifikasi, dan mengklasifikasikan opini atau emosi yang terdapat dalam suatu teks sehingga dapat diketahui apakah suatu ulasan bersifat positif, negatif, atau netral [7]. Metode ini banyak digunakan dalam penelitian yang berkaitan dengan opini publik terhadap suatu produk, layanan, maupun fenomena digital yang berkembang di masyarakat.

Salah satu algoritma yang sering digunakan dalam analisis sentimen adalah Naïve Bayes Classifier. Algoritma ini merupakan metode klasifikasi berbasis probabilitas yang menggunakan prinsip Teorema Bayes dengan asumsi independensi antar fitur. Metode ini banyak digunakan dalam klasifikasi teks karena memiliki struktur komputasi yang sederhana serta mampu memproses data teks dalam jumlah besar secara efisien [8]. Selain itu, algoritma Naïve Bayes juga dikenal memiliki performa yang baik dalam melakukan klasifikasi teks seperti ulasan pengguna pada aplikasi maupun komentar yang terdapat pada media sosial [9].

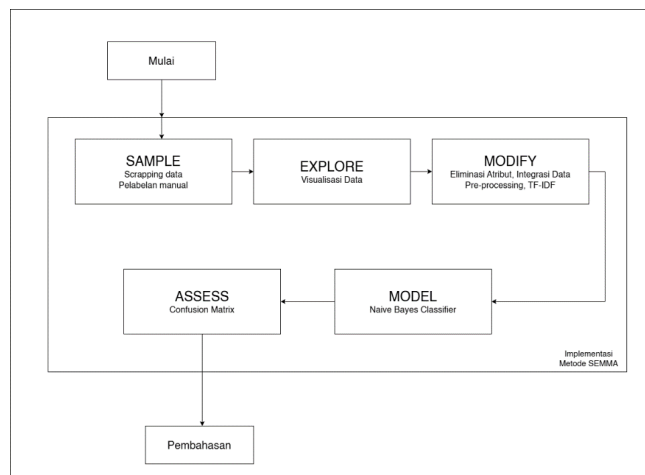
Beberapa penelitian sebelumnya juga menunjukkan bahwa algoritma Naïve Bayes memiliki tingkat akurasi yang cukup baik dalam mengklasifikasikan sentimen pengguna pada berbagai platform digital [10]. Penelitian lain juga menunjukkan bahwa metode text mining dan analisis sentimen dapat digunakan untuk memahami persepsi masyarakat terhadap layanan digital berdasarkan opini yang disampaikan oleh pengguna melalui media sosial. Dengan demikian, penerapan metode analisis sentimen menggunakan algoritma Naïve Bayes menjadi pendekatan yang relevan untuk menganalisis opini pengguna terhadap pembelian Robux pada platform resmi maupun non-resmi.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis sentimen

pengguna media sosial X terhadap pengalaman pembelian Robux menggunakan algoritma Naïve Bayes. Hasil penelitian ini diharapkan dapat memberikan gambaran mengenai persepsi pengguna terhadap transaksi pembelian Robux serta mengevaluasi performa algoritma Naïve Bayes dalam melakukan klasifikasi sentimen pada data teks media sosial [11].

METODE PENELITIAN

Metode penelitian yang digunakan dalam penelitian ini adalah pendekatan data mining untuk menganalisis sentimen pengguna media sosial X terhadap pengalaman pembelian Robux. Proses penelitian dilakukan dengan menggunakan kerangka kerja SEMMA (Sample, Explore, Modify, Model, Assess). Metode SEMMA dipilih karena mampu menggambarkan tahapan pengolahan data secara sistematis mulai dari pengambilan data hingga evaluasi model yang dihasilkan sehingga memudahkan peneliti dalam melakukan analisis data teks secara terstruktur.



Gambar 1. Rancangan Penelitian

Tahap sample merupakan tahap pengumpulan data yang digunakan dalam penelitian. Data diperoleh dari platform media sosial X berupa cuitan pengguna yang berkaitan dengan pembelian Robux. Pengumpulan data dilakukan dengan teknik scraping menggunakan beberapa kata kunci yang relevan seperti “beli robux”, “top up robux”, dan kata kunci lain yang berkaitan dengan transaksi Robux. Data yang dikumpulkan merupakan cuitan berbahasa Indonesia pada rentang waktu tertentu sehingga diperoleh dataset yang kemudian digunakan sebagai bahan analisis dalam penelitian ini.

Tahap explore bertujuan untuk memahami karakteristik data yang telah dikumpulkan. Pada tahap ini dilakukan eksplorasi terhadap data untuk mengetahui pola kata yang sering muncul serta memahami konteks percakapan pengguna terkait pengalaman pembelian Robux. Proses eksplorasi data juga membantu peneliti dalam melihat distribusi data dan mengidentifikasi kemungkinan adanya data yang tidak relevan sebelum dilakukan proses pengolahan lebih lanjut.

Tahap modify merupakan tahap pengolahan data atau preprocessing yang bertujuan untuk membersihkan data teks agar dapat digunakan dalam proses analisis. Pada tahap ini dilakukan beberapa proses pengolahan data seperti cleaning untuk menghapus karakter yang tidak diperlukan seperti tanda baca, angka, URL, dan simbol lainnya. Selain itu dilakukan tokenizing untuk memecah kalimat menjadi unit kata, stopwords removal untuk menghapus kata-kata umum yang tidak memiliki makna penting, serta stemming untuk mengubah kata menjadi bentuk dasarnya. Setelah proses preprocessing selesai, data kemudian diubah ke dalam bentuk numerik menggunakan metode Term Frequency–Inverse Document Frequency (TF-IDF) agar dapat diproses oleh algoritma klasifikasi.

Tahap model merupakan tahap pembangunan model klasifikasi menggunakan algoritma Naïve Bayes Classifier. Algoritma ini digunakan untuk mengklasifikasikan sentimen pengguna ke dalam tiga kategori yaitu sentimen positif, negatif, dan netral. Proses pemodelan dilakukan menggunakan

bahasa pemrograman Python dengan memanfaatkan beberapa library pendukung seperti Pandas, Scikit-learn, dan NLTK untuk membantu proses pengolahan dan analisis data.

Tahap assess merupakan tahap evaluasi terhadap model klasifikasi yang telah dibuat. Evaluasi dilakukan menggunakan confusion matrix untuk mengetahui performa model dalam mengklasifikasikan data. Beberapa metrik evaluasi yang digunakan antara lain akurasi, precision, dan recall. Hasil evaluasi tersebut digunakan untuk mengetahui tingkat kinerja algoritma Naïve Bayes dalam melakukan klasifikasi sentimen terhadap data cuitan pengguna terkait pembelian Robux pada media sosial X.

HASIL DAN PEMBAHASAN

Bagian ini memuat data-data hasil penelitian yang dapat disajikan dalam bentuk deskripsi, tabel, grafik, maupun gambar. Bagian dari pembahasan memaparkan hasil pengolahan data dan interpretasi hasil penelitian yang diperoleh serta mengaitkan dengan sumber rujukan yang relevan.

Bagian ini menyajikan hasil pengolahan data penelitian serta pembahasan terhadap temuan yang diperoleh. Data yang digunakan dalam penelitian ini berasal dari cuitan pengguna media sosial X yang berkaitan dengan pengalaman pembelian Robux. Pengumpulan data dilakukan dengan teknik scraping menggunakan beberapa kata kunci yang relevan seperti “beli robux”, “top up robux”, dan kata kunci lain yang berkaitan dengan transaksi Robux. Data yang diperoleh kemudian diseleksi sehingga hanya cuitan yang relevan dengan topik penelitian yang digunakan dalam proses analisis. Setelah proses seleksi dilakukan, diperoleh dataset yang selanjutnya digunakan dalam proses analisis sentimen.

Tahap awal yang dilakukan setelah pengumpulan data adalah proses preprocessing data teks. Tahap ini bertujuan untuk membersihkan data dari unsur-unsur yang tidak diperlukan sehingga data menjadi lebih terstruktur dan siap untuk dianalisis. Proses preprocessing meliputi beberapa tahapan yaitu *cleaning*, *tokenizing*, *stopword removal*, dan *stemming*. *Cleaning* dilakukan dengan menghapus karakter yang tidak diperlukan seperti tanda baca, angka, URL, serta simbol lain yang tidak memiliki pengaruh terhadap analisis sentimen. *Tokenizing* dilakukan dengan memecah kalimat menjadi unit kata sehingga memudahkan proses analisis teks. *Stopword removal* bertujuan untuk menghapus kata-kata umum yang sering muncul namun tidak memiliki makna penting dalam analisis seperti kata penghubung atau kata umum lainnya. Sementara itu, proses *stemming* dilakukan untuk mengubah kata ke dalam bentuk dasarnya sehingga variasi kata dengan makna yang sama dapat disederhanakan.

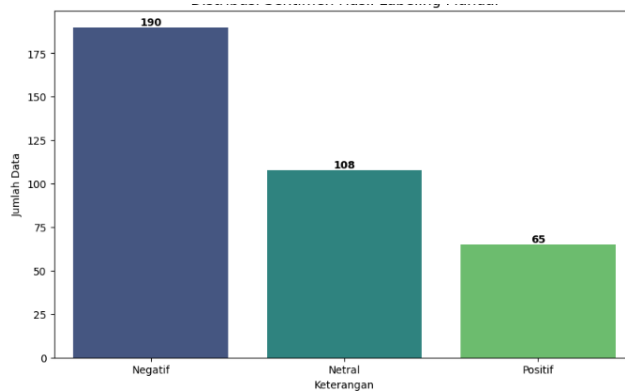
	A	B	C	D	E	F	G	H	I	J	K	L	M
	conversation	created_at	favorite_count	full_text	id_str	image_url	in_reply_to_squote_count	reply_count	retweet_count	tweet_url	user_id_str	username	
1	20189077320	Wed Feb 04 0	0	semalem aku	20189077320	https://pbs.twimg.com/media/189077320.jpg	0	0	0	https://x.com/189077320	1991734279672590338		
2	20175594884	Sat Jan 31 11	14	HELP RT PER	20175594884	https://pbs.twimg.com/media/175594884.jpg	0	13	51	https://x.com/175594884	1504067910301806594		
3	20188775191	Wed Feb 04 0	0	kalian kl mau	20188775191	https://pbs.twimg.com/media/188775191.jpg	0	1	0	https://x.com/188775191	1620788760723038215		
4	19561952368	Fri Aug 15 03:	2564	top 3 pengelu	19561952368	https://pbs.twimg.com/media/19561952368.jpg	1372249	87	8	923	https://x.com/19561952368	979395767554121729	
5	189090837193	Tue Dec 17 11	9742	got my month	189090837193	https://pbs.twimg.com/media/189090837193.jpg		108	63	561	https://x.com/189090837193	177568759471652864	
6	18850690309	Thu Jan 30 20	12749	----- Como	18850690309	https://pbs.twimg.com/media/18850690309.jpg		158	12	638	https://x.com/18850690309	1835806374749437952	
7	19572592856	Mon Aug 18 0	11532	top up 80 rob	19572592856	https://pbs.twimg.com/media/19572592856.jpg		244	169	260	https://x.com/19572592856	1210910514009628672	
8	19545093430	Sun Aug 10 11	11201	gua ngambek	19545093430	https://pbs.twimg.com/media/19545093430.jpg		420	546	550	https://x.com/19545093430	1331650559518990336	
9	20146137546	Fri Jan 23 08:	47	Hati hati peni	20146137546	https://pbs.twimg.com/media/20146137546.jpg		3	9	102	https://x.com/20146137546	1659391793472090118	
10	16459244387	Tue Apr 11 22	356	robux commis	16459244387	https://pbs.twimg.com/media/16459244387.jpg		1	51	35	https://x.com/16459244387	1258111288107884546	
11	19624118276	Mon Sep 01 0	652	DISKON ROB	19624118276	https://pbs.twimg.com/media/19624118276.jpg		23	143	31	https://x.com/19624118276	1660972969571803137	
12	18292121620	Thu Aug 29 17	306	Top creators	18292121620	https://pbs.twimg.com/media/18292121620.jpg		2	26	14	https://x.com/18292121620	950000258695815169	
13	20176390888	Sat Jan 31 16	2	Scammer!	20176390888	https://pbs.twimg.com/media/20176390888.jpg		3	1	7	https://x.com/20176390888	1438034162422714368	
14	18094828286	Sat Jul 06 07:	2013	roblox used	18094828286	https://pbs.twimg.com/media/18094828286.jpg		2	26	220	https://x.com/18094828286	2820489420	
15	19562591237	Fri Aug 15 07:	347	200 robux gw	19562591237	https://pbs.twimg.com/media/19562591237.jpg	123935	3	285	251	https://x.com/19562591237	1377403381392797698	
16	18081664186	Tue Jul 02 15:	5846	i found a way	18081664186	https://pbs.twimg.com/media/18081664186.jpg		289	86	282	https://x.com/18081664186	1637449618480054275	
17	19093327864	Mon Apr 07 15:	438	105 robux to	19093327864	https://pbs.twimg.com/media/19093327864.jpg		1	14	32	https://x.com/19093327864	1840831435722407940	
18	20178128288	Sun Feb 01 04	1996	kode promo T	20178128288	https://pbs.twimg.com/media/20178128288.jpg		48	95	94	https://x.com/20178128288	1247963289226588160	
19	20183539059	Mon Feb 02 11	2	ak abis ngasi	20183539059	https://pbs.twimg.com/media/20183539059.jpg		0	4	0	https://x.com/20183539059	1685835968287588352	
20	16285310145	Wed Feb 22 2	308	robux commis	16285310145	https://pbs.twimg.com/media/16285310145.jpg		0	37	37	https://x.com/16285310145	1258111288107884546	
21	18974485923	Thu Mar 06 01	900	@RBXevents	18974485923	https://pbs.twimg.com/media/18974485923.jpg		1	19	14	https://x.com/18974485923	1807681829513863168	

Gambar 2. Hasil Scrapping

Setelah proses *preprocessing* selesai dilakukan, tahap berikutnya adalah transformasi data teks menjadi data numerik menggunakan metode *Term Frequency-Inverse Document Frequency* (TF-IDF). Metode ini digunakan untuk memberikan bobot pada setiap kata yang muncul dalam dokumen berdasarkan frekuensi kemunculan kata tersebut dalam dokumen dan tingkat

kemunculannya pada seluruh dokumen. Dengan menggunakan metode TF-IDF, kata yang memiliki tingkat kemunculan tinggi pada suatu dokumen tetapi jarang muncul pada dokumen lain akan memiliki bobot yang lebih tinggi sehingga dianggap lebih penting dalam proses klasifikasi. Proses ini diperlukan karena algoritma klasifikasi tidak dapat memproses data dalam bentuk teks secara langsung.

Setelah data diubah menjadi bentuk numerik, tahap selanjutnya adalah proses pelabelan sentimen terhadap data yang telah dikumpulkan. Pelabelan dilakukan untuk mengelompokkan data ke dalam tiga kategori sentimen yaitu sentimen positif, sentimen negatif, dan sentimen netral. Proses pelabelan ini bertujuan untuk memberikan label kelas pada setiap data sehingga dapat digunakan dalam proses pelatihan model klasifikasi. Distribusi dan visualisasi jumlah data pada masing-masing kategori sentimen dapat dilihat pada gambar di bawah.



Gambar 3. Distribusi Visualisasi

Berdasarkan hasil distribusi data sentimen yang diperoleh, terlihat bahwa jumlah data pada masing-masing kategori memiliki perbedaan tertentu. Perbedaan jumlah data pada setiap kategori sentimen dapat mempengaruhi proses pembelajaran model klasifikasi karena model akan belajar berdasarkan pola yang terdapat pada data yang tersedia. Oleh karena itu, analisis distribusi data menjadi penting untuk memahami karakteristik dataset sebelum dilakukan proses pemodelan.

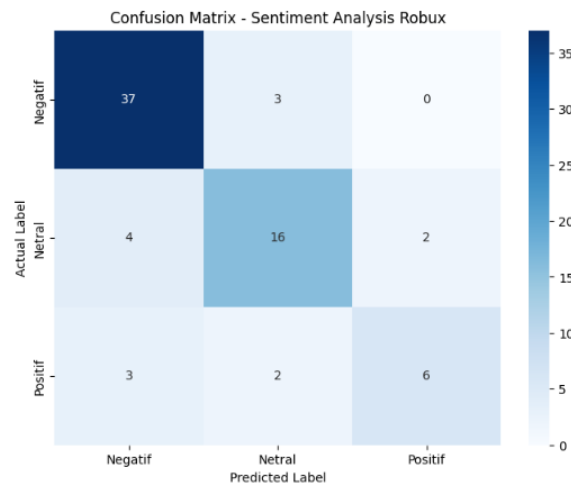
Tahap selanjutnya adalah pembangunan model klasifikasi menggunakan algoritma Naïve Bayes Classifier. Algoritma ini merupakan metode klasifikasi berbasis probabilitas yang menggunakan prinsip Teorema Bayes dengan asumsi independensi antar fitur. Algoritma Naïve Bayes banyak digunakan dalam analisis teks karena memiliki struktur komputasi yang sederhana serta mampu menghasilkan performa yang baik dalam klasifikasi dokumen teks. Pada penelitian ini, dataset yang telah diproses kemudian dibagi menjadi dua bagian yaitu data latih (*training data*) dan data uji (*testing data*). Data latih digunakan untuk melatih model agar dapat mengenali pola kata yang berkaitan dengan masing-masing kategori sentimen, sedangkan data uji digunakan untuk mengevaluasi kemampuan model dalam mengklasifikasikan data yang belum pernah dipelajari sebelumnya.

Proses pemodelan dilakukan menggunakan bahasa pemrograman Python dengan memanfaatkan beberapa *library* pendukung seperti Pandas, Scikit-learn, dan NLTK. Library tersebut digunakan untuk membantu proses pengolahan data, ekstraksi fitur, serta pembangunan model klasifikasi. Setelah model selesai dilatih menggunakan data latih, model kemudian diuji menggunakan data uji untuk melihat sejauh mana model mampu mengklasifikasikan data dengan benar.

Hasil dari proses klasifikasi kemudian dianalisis menggunakan *confusion matrix* untuk mengevaluasi performa model yang telah dibangun. *Confusion matrix* merupakan metode evaluasi yang digunakan untuk membandingkan hasil prediksi model dengan label sebenarnya dari data. Melalui confusion matrix dapat diketahui jumlah data yang diklasifikasikan dengan benar maupun

yang mengalami kesalahan klasifikasi.

Berdasarkan hasil evaluasi menggunakan *confusion matrix*, kemudian dihitung beberapa metrik evaluasi seperti akurasi, precision, dan recall untuk mengetahui performa model secara lebih rinci. Akurasi menunjukkan tingkat ketepatan model dalam mengklasifikasikan seluruh data yang diuji. Precision menunjukkan tingkat ketepatan prediksi model terhadap data yang diprediksi sebagai kelas tertentu, sedangkan recall menunjukkan kemampuan model dalam menemukan seluruh data yang termasuk dalam suatu kelas.



Gambar 4. Klasifikasi Naïve Bayes dengan tiga skenario

Hasil analisis sentimen menunjukkan bahwa opini pengguna media sosial X terhadap pembelian Robux memiliki kecenderungan yang beragam. Sebagian pengguna menunjukkan sentimen positif yang umumnya berkaitan dengan kemudahan dalam melakukan pembelian Robux serta ketersediaan berbagai metode pembayaran yang memudahkan pengguna. Di sisi lain, terdapat pula sentimen negatif yang biasanya berkaitan dengan harga yang dianggap mahal, kekhawatiran terhadap keamanan transaksi, serta pengalaman pengguna yang kurang memuaskan ketika melakukan pembelian melalui pihak ketiga. Selain itu, terdapat pula sentimen netral yang biasanya berupa pernyataan informatif atau pertanyaan pengguna terkait proses pembelian Robux.

Hasil penelitian ini menunjukkan bahwa analisis sentimen terhadap data media sosial dapat memberikan gambaran mengenai persepsi pengguna terhadap suatu layanan digital. Temuan ini juga menunjukkan bahwa metode *text mining* dan algoritma Naïve Bayes dapat digunakan secara efektif untuk mengklasifikasikan opini pengguna berdasarkan data teks yang berasal dari media sosial. Dengan demikian, penelitian ini dapat memberikan kontribusi dalam memahami perilaku serta persepsi pengguna dalam transaksi digital yang terjadi dalam lingkungan ekonomi virtual.

SIMPULAN

Penelitian ini bertujuan untuk menganalisis sentimen pengguna media sosial X terhadap pengalaman pembelian Robux dengan menggunakan algoritma Naïve Bayes. Berdasarkan hasil pengolahan data melalui tahapan SEMMA yang meliputi sample, explore, modify, model, dan assess, penelitian ini berhasil mengklasifikasikan opini pengguna ke dalam tiga kategori sentimen yaitu positif, negatif, dan netral. Hasil analisis menunjukkan bahwa sebagian pengguna memberikan sentimen positif terkait kemudahan proses pembelian dan ketersediaan metode pembayaran, sementara sentimen negatif umumnya berkaitan dengan harga yang dianggap relatif mahal serta kekhawatiran terhadap keamanan transaksi, terutama pada pembelian melalui pihak ketiga. Hasil evaluasi model menggunakan confusion matrix serta metrik akurasi, precision, dan recall menunjukkan bahwa algoritma Naïve Bayes memiliki kemampuan yang cukup baik dalam mengklasifikasikan sentimen pada data teks media sosial sehingga dapat digunakan sebagai metode

yang efektif dalam analisis sentimen.

Berdasarkan hasil penelitian yang telah dilakukan, disarankan agar penelitian selanjutnya dapat menggunakan jumlah data yang lebih besar serta memanfaatkan sumber data dari berbagai platform media sosial lainnya agar hasil analisis menjadi lebih representatif. Selain itu, penelitian selanjutnya juga dapat mencoba menggunakan algoritma klasifikasi lain atau mengombinasikan beberapa metode machine learning untuk membandingkan tingkat performa yang dihasilkan sehingga dapat diperoleh model klasifikasi sentimen yang lebih optimal.

DAFTAR PUSTAKA

- [1] W. M. S. Hutaglung, Tony, and N. J. Perdana, "Analisis Sentimen Pada Opini Kenaikan Harga Bahan Bakar Minyak Pada Media Sosial Twitter," *J. Sist. Inf. DAN Tek. Komput.*, vol. 8, no. 2, pp. 280–284, 2023.
- [2] W. A. Mohammad *et al.*, "A Survey of Data Mining Activities in Distributed Systems," *Asian J. Res. Comput. Sci.*, vol. 11, no. 4, pp. 1–18, 2021, doi: 10.9734/AJRCOS/2021/v11i430267.
- [3] T. Hidayat, "Data Mining untuk Meningkatkan Efisiensi dan Prediksi Produk Garmen Menggunakan Algoritma K-Nearest Neighbor di PT Mas Silueta," *J. KRIDATAMA SAINS DAN Teknol.*, vol. 06, no. 1, pp. 160–173, 2024.
- [4] L. Lubalu *et al.*, "Perlindungan Konsumen Terhadap Pembelian Item Digital dalam Aplikasi Game Online di Indonesia," *J. Konstr. Huk.*, vol. 3, no. 1, pp. 212–216, 2022, doi: <https://doi.org/10.22225/jkh.3.1.4464.212-216>.
- [5] U. Kulsum, M. Jajuli, and N. Sulistiyowati, "Analisis Sentimen Aplikasi WETV di Google Play Store Menggunakan Algoritma Support Vector Machine," *J. Appl. Informatics Comput.*, vol. 6, no. 2, pp. 205–212, 2022.
- [6] A. Mehrotra and R. Agarwal, "A Review of Use of Data Mining during COVID-19 Pandemic," *Turkish J. Comput. Math. Educ.*, vol. 12, no. 6, pp. 4547–4552, 2021.
- [7] Y. Findawati and M. A. Rosid, *Buku Ajar Text Mining*. 2020.
- [8] R. Andreas and D. Arymami, "Hasrat Konsumsi Virtual dalam Permainan Daring Mobile legends: Perspektif Deleuze dan Guattar," *J. Pemikir. Sociol. Univ. Gadjah Mada*, vol. 8, pp. 209–242, 2022.
- [9] M. Afdal and L. R. Elita, "Penerapan Text Mining Pada Aplikasi Tokopedia Menggunakan Algoritma K-Nearest Neighbor," *J. Ilm. Rekayasa dan Manaj. Sist. Inf.*, vol. 8, no. 1, pp. 78–87, 2022.
- [10] N. Arifin, U. Enri, and N. Sulistiyowati, "Penerapan Algoritma Support Vector Machine (Svm) Dengan Tf-Idf N-Gram Untuk Text Classification," *STRING (Satuan Tulisan Ris. dan Inov. Teknol.*, vol. 6, no. 2, 2021.
- [11] F. Alghifari and D. Juardi, "Penerapan Data Mining Pada Penjualan Makanan dan Minuman Menggunakan Metode Algoritma Naïve Bayes," *J. Ilm. Inform.*, vol. 2337, no. 8379, 2021.