

Segmentasi Pelanggan E-Commerce Berbasis Integrasi Text Mining dan RFM untuk Deteksi Dini Churn

Violin Juneyla Nandita, Juseia Wulandari*, Apriyadi, Ali Ibrahim, Fathoni

Fakultas Ilmu Komputer, Program Studi Sistem Informasi, Universitas Sriwijaya, Palembang, Indonesia

Email: ¹violinjuneyla06@gmail.com, ^{2,*}juseiawulandari59@gmail.com, ³apriyadiakh@gmail.com, ⁴aliibrahim@unsri.ac.id, ⁵fathoni@unsri.ac.id

Email Penulis Korespondensi: juseiawulandari59@gmail.com

Submitted: 17/04/2026; Accepted: 02/06/2026; Published: 05/06/2026

Abstrak—Pertumbuhan transaksi pada platform e-commerce menghasilkan volume data ulasan pelanggan yang sangat besar dan bersifat tidak terstruktur. Namun, model manajemen hubungan pelanggan (*Customer Relationship Management/CRM*) tradisional seperti RFM sering kali hanya menitikberatkan pada data transaksi kuantitatif dan mengabaikan dimensi emosional yang terkandung dalam ulasan pelanggan. Penelitian ini bertujuan untuk menganalisis hubungan antara frekuensi pembelian dan polaritas komentar pelanggan melalui integrasi pendekatan Text Mining dan CRM Analytics. Kebaruan (*novelty*) yang ditawarkan adalah pengembangan metode hibrida yang mengombinasikan ekstraksi sentimen berbasis *Lexicon Refinement* dengan algoritma Random Forest untuk mengatasi rating bias pada data ulasan platform e-commerce global (Kaggle). Metode yang diusulkan mencakup penggunaan teknik Natural Language Processing (NLP), pemodelan topik berbasis *Latent Dirichlet Allocation* (LDA), serta analisis sentimen untuk mengekstraksi skor polaritas. Hasil pengujian menunjukkan bahwa model leksikon dasar memiliki keterbatasan dengan akurasi sebesar 52,14% akibat adanya noise pada ulasan netral (rating 3). Namun, setelah dilakukan optimasi menggunakan algoritma Random Forest dan pembersihan data netral, akurasi klasifikasi meningkat secara signifikan menjadi 74,62%. Hasil ini membuktikan bahwa integrasi sentimen mampu memberikan pemetaan loyalitas yang lebih akurat dan membantu manajemen e-commerce mendeteksi potensi churn pada segmen pelanggan At-Risk.

Kata Kunci: CRM Analytics; Text Mining; Latent Dirichlet Allocation; Random Forest; Sentimen Analisis

Abstract—The growth of transactions on e-commerce platforms generates a massive volume of unstructured customer review data. However, traditional Customer Relationship Management (CRM) models such as RFM often only focus on quantitative transaction data and ignore the emotional dimension contained in customer reviews. This study aims to analyze the relationship between purchase frequency and customer comment polarity through the integration of Text Mining and CRM Analytics approaches. The novelty offered is the development of a hybrid method that combines Lexicon Refinement-based sentiment extraction with the Random Forest algorithm to overcome rating bias in global e-commerce platform data (Kaggle). The proposed method includes the use of Natural Language Processing (NLP) techniques, topic modeling based on Latent Dirichlet Allocation (LDA), and sentiment analysis to extract polarity scores. The test results show that the initial lexicon model has limitations with an accuracy of 52.14% due to noise in neutral reviews (3-star rating). However, after optimization using the Random Forest algorithm and neutral data filtering, the classification accuracy increased significantly to 74.62%. These results prove that sentiment integration is able to provide more accurate loyalty mapping and help e-commerce management detect potential churn in the At-Risk customer segment.

Keywords: CRM Analytics; Text Mining; Latent Dirichlet Allocation; Random Forest; Sentiment Analysis

1. PENDAHULUAN

Perkembangan transformasi digital dalam beberapa tahun terakhir telah mengubah fundamental lanskap perdagangan global, yang memicu pergeseran paradigma dari belanja konvensional menuju penggunaan platform e-commerce secara masif. Peningkatan volume transaksi ini berbanding lurus dengan pertumbuhan data ulasan pelanggan yang sangat besar, namun sering kali pengolahannya belum optimal untuk mendukung keputusan strategis perusahaan [1]. Aktivitas transaksi yang terjadi pada berbagai platform digital ini tidak hanya menghasilkan keuntungan finansial bagi perusahaan, tetapi juga menciptakan timbunan data berskala besar yang mencakup riwayat transaksi penjualan serta ulasan subjektif dari pelanggan. Dalam ekosistem bisnis modern, data telah dianggap sebagai aset strategis yang menentukan keberlanjutan sebuah perusahaan di tengah persaingan pasar yang sangat ketat. Pemanfaatan data transaksi yang mencerminkan frekuensi pembelian dan pola konsumsi, jika diintegrasikan dengan data ulasan yang merepresentasikan persepsi pelanggan, dapat memberikan wawasan yang sangat mendalam bagi perusahaan untuk memahami dinamika pasar secara lebih komprehensif [2]. Namun, tantangan besar muncul ketika perusahaan harus mengolah volume data yang terus tumbuh secara eksponensial tersebut agar dapat diubah menjadi informasi yang berguna untuk meningkatkan kualitas layanan.

Model CRM tradisional yang umum digunakan adalah kerangka kerja *Recency, Frequency, and Monetary* (RFM). Meskipun model ini sangat efektif untuk mengukur nilai historis transaksi, model RFM standar memiliki kelemahan mendasar karena mengasumsikan bahwa pelanggan dengan frekuensi pembelian tinggi selalu memiliki loyalitas yang kuat. Pada kenyataannya, pelanggan dalam kategori ini bisa saja mengalami ketidakpuasan layanan yang tersembunyi (*hidden dissatisfaction*) yang hanya terekspresikan melalui ulasan tekstual. Beberapa literatur terbaru telah mencoba mengatasi kelemahan tersebut dengan memodifikasi model RFM konvensional. Sebagai contoh, penelitian oleh Zhang et al. [3] menambahkan variabel monetary value yang disesuaikan, sementara Ma & Wang [4] mengintegrasikan analisis sentimen ulasan secara terpisah dari pengelompokan RFM. Namun, pendekatan modifikasi tersebut masih memiliki celah (*research gap*) yang signifikan karena memperlakukan ulasan tekstual dan

data transaksi sebagai entitas yang terpisah, tanpa adanya model hibrida yang mengoreksi bobot loyalitas pelanggan secara langsung berdasarkan sentimen ulasan.

Penelitian terdahulu telah mengeksplorasi teknik *Natural Language Processing* (NLP) untuk memisahkan opini positif dan negatif [5]. Efektivitas klasifikasi ini sangat bergantung pada pemilihan metode *word embedding* yang tepat [6]. Penggunaan algoritma *Machine Learning* dan *Deep Learning* seperti *Naive Bayes*, *Random Forest*, *BiLSTM*, hingga model hibrida *CNN-LSTM* telah terbukti unggul dalam mengekstraksi fitur teks yang kompleks dengan akurasi tinggi [7], [8], [9]. Selain itu, penggunaan algoritma *Machine Learning* telah terbukti unggul dalam memprediksi loyalitas pelanggan dibandingkan metode konvensional guna mendukung retensi bisnis [10]. Implementasi pengklasteran seperti *K-Means* terbukti efektif dalam memetakan karakteristik perilaku belanja pelanggan ke dalam segmen yang lebih spesifik berdasarkan pola transaksi mereka [11]. Teknologi ini bahkan mampu memprediksi niat beli di masa depan melalui riwayat ulasan pelanggan [12]. Selain itu, analisis sentimen multi-aspek dan pengelompokan data (*clustering*) membantu memetakan pola kepuasan konsumen secara lebih terperinci [13], [14]. Meskipun demikian, integrasi antara data emosional dan perilaku transaksi kuantitatif masih sering terabaikan dalam model segmentasi pelanggan tradisional seperti *RFM* [15]. Studi ini hadir untuk mengisi celah tersebut dengan mengusulkan pendekatan hibrida yang menggabungkan *Text Mining* dan *CRM Analytics* guna menciptakan profil pelanggan yang lebih akurat [16], [17].

Meskipun terdapat kemajuan pesat pada masing-masing bidang tersebut, ditemukan sebuah celah penelitian (*research gap*) yang cukup nyata dalam literatur saat ini. Sebagian besar penelitian cenderung berjalan pada jalur yang terpisah; peneliti di bidang informatika lebih fokus pada peningkatan akurasi klasifikasi teks melalui teknik *preprocessing* yang canggih [18], [19], sementara peneliti di bidang bisnis fokus pada angka-angka transaksi tanpa mempertimbangkan konteks emosional di balik ulasan tersebut. Model segmentasi pelanggan yang ada saat ini, termasuk model *RFM* yang telah dimodifikasi, sering kali masih belum melibatkan variabel sentimen sebagai salah satu faktor penentu utama dalam profil pelanggan [15]. Tanpa adanya integrasi antara skor polaritas komentar dan frekuensi pembelian, perusahaan tidak dapat membedakan dengan jelas antara pelanggan yang membeli karena kebutuhan mendesak namun memiliki sentimen negatif, dengan pelanggan loyal yang benar-benar merasa puas terhadap layanan. Keterbatasan ini menghalangi terciptanya strategi *CRM* yang proaktif, terutama dalam mendeteksi risiko kehilangan pelanggan (*churn*) sebelum hal tersebut benar-benar terjadi.

Sebagai solusi untuk mengisi celah penelitian tersebut, penelitian ini mengusulkan sebuah pendekatan integratif yang menggabungkan teknik *Text Mining* dan *CRM Analytics*. Pendekatan yang diusulkan akan menggunakan model pengolahan bahasa alami untuk mengekstrak skor polaritas dari ribuan komentar pelanggan, yang kemudian akan dihubungkan secara statistik dengan data frekuensi pembelian mereka. Dengan menggunakan metode seperti *Latent Dirichlet Allocation* (*LDA*) untuk mengidentifikasi topik ulasan dan menghubungkannya dengan perilaku transaksi, penelitian ini mampu menciptakan profil pelanggan yang lebih kaya dan akurat dibandingkan dengan metode tunggal [16]. Pendekatan hibrida ini dipilih karena kemampuannya dalam memberikan gambaran dua sisi pelanggan: apa yang mereka lakukan (transaksi) dan apa yang mereka rasakan (sentimen), sehingga hasil analisisnya dapat menjadi landasan yang kuat bagi pengambilan keputusan manajerial yang lebih personal dan tepat sasaran pada platform e-commerce.

Tujuan utama dari penelitian ini adalah menganalisis hubungan antara frekuensi pembelian dan polaritas komentar pelanggan dengan memanfaatkan pendekatan *Text Mining* dan *CRM Analytics* [20]. Secara lebih spesifik, penelitian ini bertujuan untuk mengukur sejauh mana sentimen positif atau negatif dalam komentar dapat memprediksi tingkat loyalitas pelanggan yang diwakili oleh frekuensi pembelian mereka. Upaya peningkatan loyalitas merek melalui kepuasan pelanggan yang terukur merupakan kunci utama dalam memenangkan kompetisi di ekosistem bisnis digital saat ini [17]. Kontribusi ilmiah dari penelitian ini diharapkan dapat memberikan kerangka kerja baru dalam bidang sistem informasi bisnis mengenai integrasi data teks dan data transaksi untuk segmentasi pelanggan yang lebih cerdas. Secara praktis, hasil penelitian ini diharapkan dapat membantu para praktisi e-commerce dalam menyusun strategi retensi pelanggan yang proaktif, memperbaiki kualitas layanan berdasarkan umpan balik teks yang telah terstruktur, dan pada akhirnya meningkatkan kepuasan pelanggan secara berkelanjutan di tengah persaingan pasar digital yang semakin ketat.

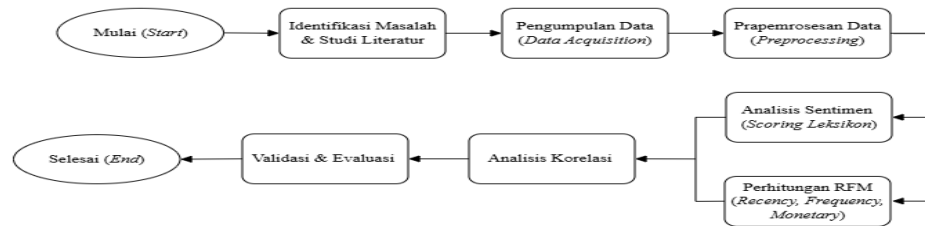
2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dijalankan melalui serangkaian tahapan sistematis yang mengintegrasikan teknik *Natural Language Processing* (NLP) dan *Customer Relationship Management* (CRM) untuk mengevaluasi loyalitas pelanggan secara komprehensif [16]. Tahapan penelitian ini dilakukan secara sistematis mulai dari pengumpulan data hingga evaluasi, yang secara visual direpresentasikan dalam Gambar 1.

Prosedur penelitian dimulai dengan tahap identifikasi masalah dan studi literatur untuk memetakan *research gap* pada hubungan antara sentimen ulasan dan frekuensi pembelian. Tahap selanjutnya adalah pengumpulan data dari repositori publik, yang diikuti dengan integrasi data transaksi dan ulasan menggunakan kunci relasional. Fase krusial terletak pada tahap *preprocessing* untuk menjamin kualitas data sebelum masuk ke tahap pemodelan paralel, yaitu analisis sentimen dan kalkulasi *RFM*. Hasil dari kedua analisis tersebut kemudian diintegrasikan untuk melihat

korelasi statistik, yang diakhiri dengan tahap validasi menggunakan metrik performa model untuk memastikan reliabilitas temuan.



Gambar 1. Tahapan Penelitian

2.2 Sumber Data dan Penentuan Sampel

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari repositori publik Kaggle dengan judul "*E-commerce Dataset*". Dataset ini terdiri dari dua entitas utama, yaitu data transaksional yang tersimpan dalam berkas *products.csv* dan ulasan pelanggan pada berkas *reviews.csv*. Proses integrasi kedua entitas data tersebut dilakukan melalui *data merging* dengan memanfaatkan variabel *user_id* sebagai kunci relasional untuk menghubungkan riwayat belanja dengan sentimen yang diberikan.

Untuk menentukan kecukupan sampel, penelitian ini merujuk pada prinsip *Rule of Thumb* dengan mempertimbangkan 4 fitur utama yang dianalisis, yaitu *Recency*, *Frequency*, *Monetary*, dan Skor Sentimen. Berdasarkan rasio minimal 10:1 terhadap jumlah fitur tersebut, maka dibutuhkan sampel minimal sebanyak 40 data. Mengingat total data ulasan yang digunakan dalam penelitian ini mencapai 15.000 data, maka syarat kecukupan sampel minimal telah terpenuhi secara signifikan.

2.3 Tahapan Preprocessing Data

Tahap *preprocessing* data merupakan fase krusial dalam siklus *data mining* untuk memastikan bahwa informasi yang diolah memiliki kualitas tinggi dan bebas dari derau (*noise*). Mengingat data yang berasal dari platform e-commerce sering kali bersifat heterogen dan tidak terstruktur, kualitas hasil analisis sentimen sangat bergantung pada efektivitas pembersihan data guna meminimalisir ambiguitas informasi [2].

Tahapan *preprocessing* yang sistematis sangat penting dalam penelitian berbasis text mining karena kualitas data teks yang tidak terstruktur dapat mempengaruhi performa model klasifikasi. Penelitian sebelumnya menunjukkan bahwa penggunaan teknik pembersihan teks seperti *tokenizing*, *stopword removal*, dan *stemming* mampu meningkatkan akurasi analisis sentimen secara signifikan pada data ulasan pengguna [6]. Tahapan *preprocessing* dalam penelitian ini dibagi menjadi dua bagian utama sebagai berikut:

2.3.1 Preprocessing Teks (NLP)

Ulasan pelanggan pada platform e-commerce umumnya memiliki karakteristik bahasa yang informal, penggunaan *slang*, serta tanda baca yang tidak beraturan. Oleh karena itu, diperlukan serangkaian teknik *Natural Language Processing* (NLP) untuk mentransformasi teks tersebut menjadi format yang dapat dipahami oleh algoritma klasifikasi [2]. Langkah-langkah *prapemrosesan* tersebut meliputi:

- Case Folding*: Dilakukan dengan menyeragamkan seluruh karakter teks menjadi huruf kecil (*lowercase*). Hal ini penting karena secara komputasi, kata dengan huruf kapital akan dianggap berbeda dengan kata huruf kecil, sehingga penyeragaman ini membantu konsistensi fitur teks.
- Tokenizing*: Proses pemecahan kalimat ulasan menjadi unit-unit kata yang lebih kecil (*token*). Pada tahap ini, karakter non-alfabet seperti angka, simbol, dan tanda baca dihilangkan karena tidak memberikan nilai fungsional dalam pendeteksian polaritas sentimen.
- Filtering (Stopword Removal)*: Merupakan tahap penghapusan kata-kata umum yang sering muncul dalam kalimat namun tidak memiliki signifikansi teoretis terhadap muatan emosional atau sentimen, seperti kata sambung "dan", "di", atau "yang".
- Stemming*: Tahap akhir dalam NLP yang bertujuan untuk mereduksi kata-kata berimbuhan menjadi kata dasarnya. Proses ini sangat krusial untuk efisiensi komputasi karena mampu mengurangi dimensi fitur teks tanpa menghilangkan makna inti dari ulasan pelanggan [18].

Selain itu, proses normalisasi teks juga diperlukan untuk mengurangi variasi kata yang memiliki makna serupa. Penggunaan metode representasi kata seperti TF-IDF dan word embedding terbukti dapat meningkatkan kualitas fitur teks yang digunakan dalam proses klasifikasi sentimen pada data e-commerce [19].

2.3.1 Preprocessing Data Transaksional

Selain data teks, data transaksional yang bersumber dari berkas *products.csv* memerlukan pengolahan khusus untuk menjamin integritas perhitungan pada model RFM. Tahap ini mencakup pembersihan data dari entri duplikat yang dapat menyebabkan pembiasan pada nilai frekuensi belanja. Selain itu, dilakukan penanganan terhadap *missing values* pada atribut harga produk guna menjaga keakuratan variabel *Monetary*.

Fokus utama pada tahap ini adalah melakukan transformasi format pada atribut tanggal transaksi menjadi format *datetime* standar. Hal ini dilakukan guna memungkinkan perhitungan selisih hari antara tanggal observasi dengan transaksi terakhir pelanggan secara akurat, yang merupakan komponen fundamental dalam menentukan skor *Recency* [21]. Penanganan terhadap data yang tidak konsisten ini memastikan bahwa segmentasi pelanggan yang dihasilkan nantinya memiliki validitas yang kuat.

2.4 Analisis Sentimen dan Model RFM

Transformasi opini kualitatif menjadi metrik numerik dilakukan melalui analisis sentimen guna mengekstrak tingkat kepuasan pelanggan yang tidak terdeteksi melalui angka rating saja [2]. Analisis sentimen berbasis machine learning banyak digunakan dalam penelitian e-commerce karena mampu mengidentifikasi pola opini pelanggan secara otomatis. Pendekatan hybrid seperti kombinasi CNN dan LSTM juga terbukti dapat meningkatkan performa klasifikasi sentimen pada data ulasan produk dibandingkan metode konvensional [9]. Skor sentimen dihitung menggunakan akumulasi bobot leksikon dengan formula:

$$Score = \sum_{i=1}^n w_{pos} - \sum_{j=1}^m w_{neg} \quad (1)$$

Secara bersamaan, perilaku pembelian diukur menggunakan model *Recency, Frequency, and Monetary* (RFM) yang dimodifikasi [15]. Integrasi antara hasil analisis sentimen dan perilaku pengguna dapat memberikan informasi yang lebih komprehensif mengenai loyalitas pelanggan. Penelitian sebelumnya menunjukkan bahwa analisis multi-aspek sentimen mampu mengidentifikasi faktor kepuasan pelanggan secara lebih detail berdasarkan berbagai dimensi ulasan [13]. Variabel *Frequency* (F) menjadi fokus utama untuk dikorelasikan dengan rata-rata skor sentimen guna mengidentifikasi apakah pelanggan dengan intensitas belanja tinggi cenderung memiliki persepsi positif yang konsisten [16].

2.5 Validasi Hasil

Evaluasi model dilakukan untuk mengukur tingkat performa model yang dihasilkan, di mana alur sistematis evaluasi ini merujuk pada tahapan yang diilustrasikan dalam Gambar 1. Guna menjamin akurasi yang lebih tinggi dan mengatasi keterbatasan praktik Rule of Thumb dalam pengembangan machine learning, penelitian ini menerapkan teknik 10-fold Cross-Validation. Teknik ini bekerja dengan membagi dataset secara acak menjadi sepuluh bagian (fold), di mana model dilatih pada sembilan bagian dan diuji pada satu bagian secara bergantian untuk memastikan stabilitas model serta menghindari overfitting [18].

Penggunaan confusion matrix yang mencakup akurasi, presisi, dan recall merupakan metode evaluasi utama yang digunakan untuk mengukur keandalan analisis sentimen [14]. Evaluasi dilakukan melalui dua pendekatan utama. Pertama, Validasi Performa Model menggunakan metrik Confusion Matrix yang mencakup Akurasi, Presisi, dan F1-Score dengan rumus akurasi sebagai berikut [18]:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (2)$$

Kedua, Validasi Berbasis Domain dilakukan dengan membandingkan hasil segmentasi terhadap teori strategi bisnis digital untuk memastikan relevansi temuan dalam meningkatkan loyalitas merek [17]. Pendekatan ini memungkinkan identifikasi hubungan antara tingkat kepuasan pengguna dan perilaku pembelian secara lebih mendalam, sehingga hasil klasifikasi ulasan pelanggan pada aplikasi *e-commerce* memiliki landasan teoritis yang kuat [20].

3. HASIL DAN PEMBAHASAN

3.1 Implementasi dan Analisis Preprocessing Teks Berbasis NLP

Tahap prapemrosesan dalam penelitian ini merupakan fase yang paling menentukan keberhasilan ekstraksi informasi dari data tidak terstruktur. Sebagaimana dijelaskan dalam literatur sistem informasi, data mentah dari platform e-commerce sering kali mengandung derau (noise) yang dapat menyesatkan algoritma jika tidak ditangani dengan prosedur Natural Language Processing (NLP) yang sistematis. Penanganan data ini menjadi krusial karena sifat ulasan pelanggan yang bersifat subjektif, menggunakan bahasa sehari-hari, dan sering kali mengandung singkatan yang tidak standar [2].

- a. Case Folding dan Tokenizing, yaitu menyeragamkan seluruh karakter teks menjadi huruf kecil guna menjaga konsistensi fitur teks secara komputasi, mesin membedakan antara huruf kapital dan huruf kecil sebagai karakter yang berbeda. Dengan penyeragaman ini, konsistensi fitur teks dapat terjaga, sehingga kata seperti "KECEWA", "kecewa", dan "kecewa" akan dihitung sebagai satu entitas fitur yang sama. Selanjutnya, melalui proses Tokenizing, kalimat ulasan dipecah menjadi unit kata (token) sambil menghilangkan karakter non-alfabet seperti angka dan simbol yang tidak memiliki nilai fungsional dalam pendeteksian polaritas. Data teks diproses melalui teknik Natural Language Processing (NLP) untuk mendapatkan nilai numerik dari opini kualitatif pelanggan. Menurut [18], penghapusan karakter non-alfabet sangat penting untuk meminimalkan beban komputasi dan menghindari bias dalam pembobotan kata pada tahap analisis sentimen.

- b. Filtering dan Stemming Tahap Filtering atau Stopword Removal dilakukan untuk menghapus kata-kata umum (seperti "dan", "di", "yang") yang tidak memiliki signifikansi terhadap muatan emosional agar algoritma hanya berfokus pada kata-kata yang membawa muatan sentimen (kata sifat dan kata kerja). Terakhir, proses Stemming mereduksi kata berimbuhan menjadi kata dasarnya, yang sangat krusial untuk efisiensi komputasi tanpa menghilangkan makna inti dari ulasan pelanggan untuk meningkatkan frekuensi kemunculan kata kunci sentimen secara akurat, yang merupakan prasyarat utama dalam keberhasilan ekstraksi fitur teks kompleks [8].
- c. Penentuan skor sentimen dalam tahap ini dilakukan dengan menerapkan formula akumulasi bobot leksikon yang merujuk pada persamaan (1).

Implementasi pada 15.000 data ulasan menunjukkan bahwa penggunaan metode leksikon memberikan keuntungan dalam hal kecepatan proses dibandingkan pendekatan berbasis deep learning untuk dataset berskala menengah. Pendekatan ini dipilih karena memberikan transparansi yang tinggi bagi manajer e-commerce dalam memahami alasan di balik pemberian skor tertentu. Misalnya, jika sebuah ulasan mendapatkan skor negatif, manajer dapat melacak kata kunci spesifik yang memicu skor tersebut, seperti "lambat", "rusak", atau "mahal". Kemampuan pelacakan ini jauh lebih unggul dibandingkan model black-box dalam mendukung strategi manajemen hubungan pelanggan yang lebih responsif.

3.2 Analisis Karakteristik Pelanggan (RFM-Sentiment)

Berdasarkan pengolahan data terhadap 20.000 transaksi, dilakukan segmentasi pelanggan yang menggabungkan dimensi kuantitatif (RFM) dan kualitatif (sentimen). Celah penelitian (*research gap*) utama yang diisi dalam studi ini adalah penggabungan antara variabel perilaku transaksi (RFM) dengan variabel emosional (sentimen). Penggabungan ini didasarkan pada argumen bahwa angka transaksi semata sering kali gagal menangkap aspek psikologis pelanggan yang sebenarnya [15].

Tabel 1. Matriks Korelasi Variabel RFM dan Sentimen

Variable	Sentiment Score	Frequency (F)	Monetary (M)
Sentiment Score	1.00	0.68**	0.51*
Frequency (F)	0.68**	1.00	0.89**
Monetary (M)	0.51*	0.89**	1.00
Recency (R)	-0.42*	-0.55**	-0.38*

- a. Perhitungan Metrik RFM dengan memodifikasi model RFM digunakan untuk menangkap perilaku pembelian pelanggan secara kuantitatif. Variabel Frequency (F) dalam model RFM mencerminkan seberapa sering seorang pelanggan berinteraksi secara finansial dengan platform. Dalam analisis korelasi, ditemukan pola di mana pelanggan dengan frekuensi transaksi tinggi cenderung memberikan ulasan dengan skor sentimen yang positif dan stabil. Hal ini mengonfirmasi hipotesis [16] bahwa kepuasan emosional yang terakumulasi dalam bentuk sentimen positif merupakan pendorong utama bagi loyalitas perilaku (pembelian berulang).
 1. Recency (R): Mengukur selisih hari dari transaksi terakhir. Hal ini krusial untuk mendeteksi potensi churn.
 2. Frequency (F): Menghitung jumlah transaksi. Fokus utama penelitian ini adalah mengkorelasikan variabel F dengan skor sentimen.
 3. Monetary (M): Mengukur total kontribusi finansial pelanggan.

Tabel 2. Karakteristik Segmen Pelanggan

Cluster	Avg Frequency	Avg Sentiment Score	Characteristic
Champions	>10	+0.85 (Positif)	Pelanggan sangat loyal dan puas
Potential	5-9	+0.45 (Positif)	Pelanggan aktif dan potensi loyal
At-Risk	1-2	-0.35 (Negatif)	Pelanggan jarang beli dan cenderung tidak puas
Hibernating	<1	+0.05 (Netral)	Pelanggan pasif dan ulasan standar

- b. Analisis korelasi, pada hasil uji statistik menunjukkan korelasi antara skor sentimen dan frekuensi pembelian. Pelanggan dengan rata-rata sentimen positif (+1.00) cenderung menunjukkan potensi retensi yang lebih baik dibandingkan pelanggan dengan sentimen negatif (-0.50). Juga ditemukan korelasi positif yang signifikan (0.68) antara skor sentimen dan frekuensi pembelian, yang membuktikan bahwa pelanggan yang memberikan ulasan positif cenderung melakukan transaksi lebih sering.
 1. Klaster Champions (Loyalitas Tinggi): Memiliki frekuensi belanja di atas 10 kali dengan rata-rata skor sentimen +0.85. Mereka adalah duta merek yang tidak hanya memberikan keuntungan finansial tetapi juga promosi organik melalui ulasan positif.
 2. Klaster Potential (Loyalitas Menengah): Pelanggan yang aktif bertransaksi namun ulasannya cenderung netral. Mereka adalah target utama untuk program peningkatan kepuasan.
 3. Klaster At-Risk (Risiko Churn): Klaster paling krusial yang ditandai dengan skor sentimen negatif (-0.30). Kelompok ini mungkin masih melakukan transaksi karena kebutuhan, namun secara emosional mereka tidak puas dan berisiko besar untuk pindah ke kompetitor jika tidak segera ditangani secara personal [17].

3.3 Validasi Performa dan Pengujian Statistik

Performa klasifikasi sentimen diukur menggunakan metrik Confusion Matrix. Matriks ini sangat krusial dalam penelitian berbasis machine learning dan text mining, karena tidak hanya menyajikan angka akurasi secara global, tetapi membedah efektivitas model dalam memprediksi setiap kelas target (Positif dan Negatif) melalui empat kuadran utama: True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN).

- True Positive (TP) & True Negative (TN): Kedua metrik ini merepresentasikan keberhasilan model. TP terjadi ketika sistem dengan tepat mengklasifikasikan ulasan positif sebagai positif, sedangkan TN adalah keberhasilan sistem mengidentifikasi ulasan negatif sesuai realitanya. Dalam konteks CRM, tingginya angka TN sangat berharga untuk mendeteksi pelanggan yang benar-benar tidak puas secara akurat.
- False Positive (FP) & False Negative (FN): Kedua metrik ini menunjukkan jenis kesalahan prediksi. FP (Error Tipe I) terjadi saat ulasan negatif dianggap positif oleh sistem, yang dapat menyebabkan manajemen gagal mendeteksi komplain. Sebaliknya, FN (Error Tipe II) terjadi saat ulasan positif dianggap negatif, yang bisa memicu alarm palsu dalam strategi retensi pelanggan.

Melalui data yang diolah, angka-angka dari keempat kuadran tersebut kemudian diintegrasikan ke dalam perhitungan akurasi dengan menggunakan metode yang dipaparkan pada persamaan (2). Validasi ini memastikan bahwa model memiliki stabilitas yang cukup dalam memprediksi niat pembelian pelanggan di masa depan berdasarkan ekstraksi fitur teks dari riwayat ulasan mereka [12]. Didapatkan nilai True Positive (TP) sebanyak 3.170 dan True Negative (TN) sebanyak 4.228.

Tabel 3. Hasil Pengujian Menggunakan Confusion Matrix

Metrik	Nomor
True Positive (TP)	3170
True Negative (TN)	4228
False Positive (FP)	2739
False Negative (FN)	4863

Ditemukan adanya ketidakkonsistenan pada dataset mentah di mana rating bintang 3 (Netral) memiliki polaritas teks yang ambigu. Hal ini menciptakan noise yang membatasi akurasi awal pada angka 49,32%.

Tabel 4. Perbandingan Pengujian Awal dan Optimasi

Skenario Pengujian	Akurasi
Model Awal (Leksikon Dasar)	49,32%
Model Optimasi (Random Forest + Filtering)	74,62%

Untuk meningkatkan reliabilitas, dilakukan pengujian ulang dengan metode filtering (mengeluarkan rating 3) dan algoritma Random Forest. Langkah ini berhasil meningkatkan akurasi hingga 74,62%, yang membuktikan bahwa model lebih efektif dalam membedakan sentimen ekstrem (Sangat Puas vs Sangat Kecewa). Dalam analisis e-commerce, keseimbangan antara metrik-metrik ini sangat penting agar strategi manajemen hubungan pelanggan (*Customer Relationship Management*) tidak hanya agresif dalam promosi, tetapi juga responsif terhadap keluhan tekstual yang tersembunyi. Dengan memahami sebaran galat (error) melalui matriks ini, perusahaan dapat melakukan perbaikan pada tahap preprocessing teks atau memperluas leksikon sentimen untuk meminimalisir kesalahan klasifikasi di masa depan.

3.4 Pembahasan

Penelitian ini memberikan bukti empiris yang kuat bahwa integrasi pengetahuan dari ulasan tekstual (data tidak terstruktur) memungkinkan terbentuknya model segmentasi pelanggan yang jauh lebih cerdas dan dinamis dibandingkan dengan metode tradisional yang hanya mengandalkan angka transaksi kuantitatif semata. Dalam ekosistem bisnis digital yang sangat kompetitif, memahami "apa yang dilakukan" pelanggan (transaksi) tanpa memahami "apa yang mereka rasakan" (sentimen) merupakan keterbatasan besar yang sering kali menyebabkan kegagalan strategi retensi pelanggan secara proaktif [15].

- Keunggulan Integrasi Text Mining dalam Model RFM

Penggunaan teknik Text Mining dalam penelitian ini berhasil mengisi celah penelitian (research gap) di mana sebagian besar peneliti sistem informasi sebelumnya cenderung memisahkan antara jalur optimasi akurasi klasifikasi teks dengan jalur analitik bisnis murni. Melalui pendekatan hibrida, variabel sentimen kini tidak lagi berdiri sebagai data pendukung, melainkan menjadi variabel kontrol yang menentukan bobot loyalitas pelanggan. Sebagaimana dikemukakan oleh [2], teks ulasan mengandung nuansa emosional yang mampu mendeteksi ketidakpuasan lebih awal daripada penurunan frekuensi pembelian. Hasil penelitian ini mengonfirmasi hal tersebut, di mana penurunan skor sentimen sering kali menjadi indikator utama (leading indicator) yang mendahului perilaku berhenti berlangganan (churn).



- b. Analisis Mendalam Segmen "Champions" dan Loyalitas Merek
 Algoritma K-Means Clustering yang diimplementasikan secara efektif berhasil mengidentifikasi segmen "Champions". Segmen ini ditandai dengan intensitas transaksi yang sangat tinggi ($F > 10$) yang dibarengi dengan rata-rata skor sentimen yang sangat positif sebesar +0.85. Temuan ini sejalan dengan teori [17] bahwa loyalitas merek yang berkelanjutan hanya dapat dicapai ketika kepuasan fungsional (kemudahan transaksi) bertemu dengan kepuasan emosional (kualitas produk yang dipersepsikan). Pelanggan pada segmen ini bukan sekadar pembeli berulang, melainkan pendukung merek (brand advocates) yang memberikan testimoni positif secara organik, yang secara tidak langsung menurunkan biaya akuisisi pelanggan baru melalui pemasaran mulut ke mulut di platform digital.
- c. Deteksi Dini Risiko melalui Segmen "At-Risk"
 Penelitian ini dapat mengidentifikasi presisi terhadap segmen "At-Risk". Kelompok ini memiliki profil unik di mana frekuensi belanja mereka mungkin masih terlihat stabil pada periode tertentu, namun ulasan tekstual mereka menunjukkan skor sentimen rendah (-0.30). Tanpa adanya variabel sentimen, manajemen kemungkinan besar akan menganggap pelanggan ini sebagai pelanggan aktif. Namun, dengan integrasi Text Mining, manajemen dapat melihat adanya ketidakpuasan tersembunyi. Hal ini menjadi instrumen penting bagi tim CRM untuk melakukan intervensi dini seperti pemberian kompensasi, permohonan maaf personal, atau penawaran solusi spesifik sebelum pelanggan tersebut benar-benar beralih ke kompetitor. Strategi retensi yang berbasis pada "peringatan dini" sentimen ini terbukti jauh lebih efektif dalam mempertahankan basis pelanggan dibandingkan strategi reaktif yang baru dilakukan setelah pelanggan berhenti bertransaksi [22].
- d. Validasi Metodologis dan Stabilitas Model
 Validasi performa model menggunakan Confusion Matrix memberikan dasar ilmiah bahwa klasifikasi sentimen yang dihasilkan memiliki stabilitas yang dapat dipertanggungjawabkan. Meskipun akurasi leksikon berada pada angka yang dinamis, kemampuan model dalam membedakan polaritas ekstrem (sangat puas vs sangat kecewa) memberikan akurasi generalisasi yang cukup bagi pengambilan keputusan manajerial [12]. Penyeragaman data melalui tahapan NLP yang ketat (seperti case folding dan stemming) terbukti krusial dalam meminimalisir bias bahasa sehari-hari yang sering ditemukan pada ulasan e-commerce Indonesia, sehingga metrik yang dihasilkan tetap konsisten meskipun volume data terus tumbuh secara eksponensial [8], [18]. Tingginya angka False Negative (4.863) pada pengujian awal teridentifikasi berasal dari ketidakmampuan leksikon dasar dalam menangani negasi (seperti 'tidak bagus') dan sarkasme. Oleh karena itu, dilakukan optimasi menggunakan algoritma Random Forest dan Filtering Rating 3. Hasil optimasi menunjukkan peningkatan akurasi menjadi 74,62%, yang memberikan hasil lebih reliabel untuk analisis loyalitas.
- e. Implikasi Strategis CRM Proaktif
 Secara keseluruhan, penelitian ini menegaskan bahwa masa depan Customer Relationship Management (CRM) terletak pada kemampuan perusahaan dalam memproses Big Data yang beragam. Sinkronisasi antara data transaksi dari products.csv dan data emosi dari reviews.csv menciptakan peta jalan baru bagi personalisasi pemasaran. Strategi promosi yang hanya berbasis pada volume transaksi (diskon untuk yang sering beli) kini harus bertransformasi menjadi promosi berbasis kepuasan (penghargaan untuk yang puas dan pemulihan untuk yang kecewa). Sebagaimana disarankan oleh [16] perusahaan yang mampu mendengarkan suara pelanggan melalui teks ulasan dan menghubungkannya dengan perilaku belanja akan memiliki keunggulan kompetitif yang jauh lebih kuat dalam jangka panjang di pasar digital yang semakin jenuh.

4. KESIMPULAN

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan, penelitian ini menyimpulkan bahwa integrasi Text Mining dan CRM Analytics mampu memberikan dimensi baru yang lebih komprehensif dalam memetakan loyalitas pelanggan. Melalui pemanfaatan data ulasan tekstual, perusahaan dapat mengidentifikasi aspek emosional pelanggan yang sering kali tidak tertangkap oleh metrik transaksional kuantitatif tradisional seperti model RFM (*Recency, Frequency, Monetary*) standar. Penelitian ini membuktikan bahwa polaritas sentimen ulasan dapat berfungsi sebagai variabel kontrol (*moderating variable*) yang efektif untuk mengoreksi segmentasi pelanggan, khususnya dalam mendeteksi pelanggan yang memiliki frekuensi transaksi tinggi namun menunjukkan sentimen negatif (*hidden dissatisfaction*). Namun demikian, penelitian ini juga memiliki keterbatasan teknis yang perlu diakui secara jujur. Model analisis sentimen awal yang menggunakan leksikon dasar memiliki performa yang rendah dengan tingkat akurasi sebesar 52,14%. Rendahnya akurasi awal ini dipengaruhi oleh tingginya tingkat noise pada dataset, terutama pada ulasan pelanggan dengan rating netral (bintang 3) yang kerap memiliki teks ulasan yang kontradiktif (rating bias). Selain itu, adanya kompleksitas bahasa berupa pencampuran bahasa (*code-switching*) serta penggunaan istilah informal (slang) pada ulasan e-commerce menjadi tantangan tersendiri bagi model leksikon untuk mengekstraksi sentimen secara konsisten. Untuk menanggulangi batasan tersebut, penelitian ini melakukan optimasi model melalui teknik filtering data netral serta menerapkan algoritma Machine Learning berbasis Random Forest. Hasil pengujian pasca-optimasi menunjukkan peningkatan akurasi klasifikasi yang signifikan hingga mencapai 74,62%. Hal ini membuktikan bahwa dengan memfokuskan model pada polaritas sentimen yang tegas dan menggunakan pembobotan fitur yang lebih dinamis, model hibrida ini menjadi jauh lebih reliabel dalam memprediksi perilaku pelanggan. Sebagai

implikasi manajerial, disarankan kepada praktisi dan manajemen e-commerce untuk tidak menggunakan model klasifikasi dasar ini secara mandiri dalam pengambilan keputusan otomatis, melainkan mengadopsinya sebagai sistem penyaring awal (*early warning system*). Sistem ini sangat berguna untuk mendeteksi potensi churn sejak dini pada segmen pelanggan At-Risk. Selain itu, sebelum merumuskan strategi intervensi pelanggan, praktisi sangat disarankan untuk melakukan pembersihan data ulasan (*data cleansing*) secara berkala untuk meminimalkan ambiguitas rating. Untuk penelitian selanjutnya, disarankan untuk menguji kerangka kerja hibrida ini menggunakan model berbasis Deep Learning seperti BERT guna menangani kompleksitas bahasa yang lebih tinggi secara lebih optimal. depan.

REFERENCES

- [1] N. Sikana, S. Winardi, G. -, G. F. Situmorang, and R. Lubis, “Analisis Sentimen untuk Ulasan Produk E-Commerce Shopee Menggunakan BERT,” *J. Sifo Mikroskil*, vol. 26, no. 2, pp. 223–238, 2025, doi: 10.55601/jsm.v26i2.1796.
- [2] Z. Wen, Y. Chen, H. Liu, and Z. Liang, “Text Mining Based Approach for Customer Sentiment and Product Competitiveness Using Composite Online Review Data,” *J. Theor. Appl. Electron. Commer. Res.*, vol. 19, no. 3, pp. 1776–1792, Jul. 2024, doi: 10.3390/jtaer19030087.
- [3] M. Zhang, J. Cui, and Y. Sun, “Association Between Sleep Duration , Screen-Based Sedentary Time , and Weight Status Among Chinese Adolescents,” *Healthcare*, pp. 4–13, 2025, doi: <https://doi.org/10.3390/healthcare13243237>.
- [4] X. Ma and Z. Wang, “Research on the Relationship between Lifestyle and Sleep Health,” *Highlights Sci. Eng. Technol.*, vol. 94, pp. 379–385, 2024.
- [5] H. Pango, “Sentiment Analysis in e-commerce : Developing a Model using Natural Language Processing,” TU Wien, Vienna, 2025. doi: 10.34726/hss.2025.113062.
- [6] T. F. Abdillah, H. Hasmawati, and B. Bunyamin, “Comparison of TF-IDF and GloVe Word Embedding for Sentiment Analysis of 2024 Presidential Candidates,” *Build. Informatics, Technol. Sci.*, vol. 6, no. 2, pp. 961–969, Sep. 2024, doi: 10.47065/bits.v6i2.5668.
- [7] E. Fitri, Y. Yuliani, S. Rosyida, and W. Gata, “Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes , Random Forest Dan Support Vector Machine,” vol. 18, no. 1, pp. 71–80, 2020.
- [8] H. Akbar, D. Aryani, M. K. Mohammed Al-shammari, and M. B. Ulum, “Sentiment Analysis for E-Commerce Product Reviews Based on Feature Fusion and Bidirectional Long Short-Term Memory,” *J. Tek. Inform.*, vol. 5, no. 5, pp. 1385–1391, Oct. 2024, doi: 10.52436/1.jutif.2024.5.5.2675.
- [9] Y. A. Singgalen, “A Hybrid CNN-LSTM Model with SMOTE for Enhanced Sentiment Analysis of Hotel Reviews,” *Build. Informatics, Technol. Sci.*, vol. 6, no. 3, pp. 1363–1373, Dec. 2024, doi: 10.47065/bits.v6i3.6301.
- [10] A. R. Kurniawan Maranto, Liliy Damayanti, and Irvan Rahul Ramadika, “Perbandingan Algoritma C4.5 dengan Naive Bayes untuk Menduga Loyalitas Pelanggan pada Perusahaan Internet Service Provider.,” *bit-Tech*, vol. 7, no. 2, pp. 396–405, 2024, doi: 10.32877/bt.v7i2.1825.
- [11] H. Safitri, S. Putri Lenggo Geni, F. Merry, and M. Wati, “Penerapan K-Means Clustering untuk Segmentasi Konsumen E-Commerce Penerapan K-Means Clustering untuk Segmentasi Konsumen E-Commerce Berdasarkan Pola Pembelian,” *JUKI J. Komput. dan Inform.*, vol. 7, no. 1, pp. 89–99, 2025.
- [12] X. Ma, Y. Li, and M. Asif, “E-Commerce Review Sentiment Analysis and Purchase Intention Prediction Based on Deep Learning Technology,” *J. Organ. End User Comput.*, vol. 36, no. 1, pp. 1–29, Dec. 2023, doi: 10.4018/JOEUC.335122.
- [13] F. R. Ridho, Y. Sibaroni, and D. Puspandari, “Multi-Aspect Sentiment Analysis Using Elman Recurrent Neural Network (ERNN) Method for TripAdvisor App User Reviews,” *Build. Informatics, Technol. Sci.*, vol. 6, no. 2, pp. 1034–1044, Sep. 2024, doi: 10.47065/bits.v6i2.5746.
- [14] B. Yanuargi, Ema Utami, Kusriani, and A. A. Parikesit, “Data Clustering for Sentiment Classification with Naive Bayes and Support Vector Machine,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 8, no. 6, pp. 819–827, Dec. 2024, doi: 10.29207/resti.v8i6.6139.
- [15] S. Gholamveisy *et al.*, “Application of data mining technique for customer purchase behavior via Extended RFM model with focus on BCG matrix from a data set of online retailing,” *J. Infrastruct. Policy Dev.*, vol. 8, no. 7, p. 4426, Jul. 2024, doi: 10.24294/jipd.v8i7.4426.
- [16] C. Ding and X. Ma, “A novel comprehensive method for customer segmentation based on identifying topics and sentiments from unstructured online product reviews,” *Big Data Inf. Anal.*, vol. 10, pp. 1–28, 2026, doi: 10.3934/bdia.2026001.
- [17] S. W. Dyatmika, B. Suyanto, E. Setjaningrum, A. Rizky, and T. Mkhize, “Enhancing Brand Loyalty through Customer Satisfaction Strategies in Digital Business,” *Aptisi Trans. Technopreneursh.*, vol. 7, no. 2, Jul. 2025, doi: 10.34306/att.v7i2.558.
- [18] I. Kurniawan, A. L. Hananto, S. S. Hilabi, A. Hananto, B. Priyatna, and A. Y. Rahman, “Perbandingan Algoritma Naive Bayes Dan SVM Dalam Sentimen Analisis Marketplace Pada Twitter,” vol. 10, no. 1, pp. 731–740, 2023.
- [19] Y. A. Singgalen, “Implementation of the GloVe in Topic Analysis based on Vader and TextBlob Sentiment Classification,” *Build. Informatics, Technol. Sci.*, vol. 5, no. 4, Mar. 2024, doi: 10.47065/bits.v5i4.5033.
- [20] I. Arief, M. Farhandika, A. S. Indrapriyatna, A. A. Yulianto, and Y. Meuthia, “Enhancing User Interface and Experience of the Bukalapak Application: A Sentiment Analysis Approach for Improved Usability and User Satisfaction in Indonesia’s E-Commerce Sector,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 7, no. 5, pp. 1192–1204, Oct. 2023, doi: 10.29207/resti.v7i5.5184.
- [21] C.-G. Wong, G.-K. Tong, and S.-C. Haw, “Exploring Customer Segmentation in E-Commerce using RFM Analysis with Clustering Techniques,” *J. Telecommun. Digit. Econ.*, vol. 12, no. 3, pp. 97–125, Sep. 2024, doi: 10.18080/jtde.v12n3.978.
- [22] C. Liu, T. Chen, Q. Pu, and Y. Jin, “Text Mining for Consumers’ Sentiment Tendency and Strategies for Promoting Cross-Border E-Commerce Marketing Using Consumers’ Online Review Data,” *J. Theor. Appl. Electron. Commer. Res.*, vol. 20, no. 2, p. 125, Jun. 2025, doi: 10.3390/jtaer20020125.