 ISSN:2302-8149
e-ISSN:2540-9719

Analisis Sentimen Ulasan Aplikasi *Money Lover* menggunakan *Random Forest* dan *Naïve Bayes*

Sentiment Analysis of Money Lover App Reviews using Random Forest and Naïve Bayes

[1]Nanda Aulia Salsa Bila*, [2]Wiwit Agus Triyanto, [3]Pratomo Setiaji
[1,2,3]Program Studi Sistem Informasi, Fakultas Teknik, Universitas Muria Kudus
[1,2,3]Jl. Lkr. Utara, Kayuapu Kulon, Gondangmanis, Kudus, Jawa Tengah
*e-mail: 202153016@std.umk.ac.id, at.wiwit@umk.ac.id, pratomo.setiaji@umk.ac.id

(***received*:** 18 November 2025, ***revised*:** 7 December 2025, ***accepted*:** 8 December 2025)

Abstrak

Penelitian ini bertujuan untuk menganalisis sentimen pengguna aplikasi Money Lover dan membandingkan kinerja dua algoritma *machine learning* yang berbeda, *Random Forest* dan *Naive Bayes*, dalam klasifikasi biner pada data ulasan. Sebanyak 3000 data komentar yang dikumpulkan menggunakan teknik *web scraping* yang selanjutnya diklasifikasikan ke kelas sentimen positif dan negatif. Tahap *preprocessing* terdiri dari pembersihan teks, normalisasi, tokenisasi, *stopword removal*, dan *stemming*, Tahap berikutnya proses penimbangan kata menggunakan TF-IDF untuk mengkonversi teks ke dalam bentuk vektor numerik. Hasil penelitian diperoleh gambaran mengenai kecenderungan sentimen pengguna terhadap aplikasi Money Lover sekaligus menunjukkan efektivitas kedua algoritma dalam pengolahan ulasan teks pada domain finansial. Berdasarkan evaluasi model, algoritma *Random Forest* menunjukkan performa rata – rata yang lebih unggul, dengan nilai akurasi sebesar 94%. Sementara itu algoritma *Naive Bayes* memiliki performa yang sedikit lebih rendah dengan nilai akurasi sebesar 92%. Hasil tersebut didukung dengan hasil *cross-validation* dan *ROC-Curve* yang menunjukkan konsistensi performa *Random Forest* dibandingkan *Naive Bayes*. Perbedaan performa tersebut mengindikasikan bahwa pendekatan *ensemble* seperti *Random Forest* mampu menangani variasi teks pada data ulasan dengan lebih baik sehingga memberikan hasil klasifikasi yang lebih stabil dan akurat.

Kata kunci: money lover. *naive bayes*, *random forest*, sentimen

Abstract

This study aims to analyze user sentiment toward the Money Lover application and to compare the performance of two different machine learning algorithms, Random Forest and Naïve Bayes, in binary classification of review data. A total of 3,000 comments were collected using web scraping techniques and then classified into positive and negative sentiment categories. The preprocessing stage included text cleaning, normalization, tokenization, stopword removal, and stemming. In the next stage, term weighting was performed using TF-IDF to convert the text into numerical vector representations. The results provide insights into the overall sentiment tendencies of users toward the Money Lover application and demonstrate the effectiveness of both algorithms in processing textual reviews within the financial domain. Based on model evaluation, the Random Forest algorithm achieved superior average performance, with an accuracy of 94%. Meanwhile, the Naïve Bayes algorithm showed slightly lower performance, achieving an accuracy of 92%. These findings were supported by cross-validation results and ROC curve analysis, which indicated that Random Forest consistently outperformed Naïve Bayes. The performance difference suggests that an ensemble-based approach such as Random Forest is better able to handle textual variation in review data, resulting in more stable and accurate sentiment classification.

Keywords: *money lover naive bayes, random forest, sentiment*

 ISSN:2302-8149
e-ISSN:2540-9719

1 Pendahuluan

Perkembangan teknologi modern masa kini, memengaruhi kemampuan pengelolaan keuangan. Banyak individu sering menghadapi kesulitan dalam mengontrol keuangan sehingga memengaruhi stabilitas finansial mereka[1]. Oleh karena itu, penggunaan aplikasi pengelolaan keuangan dapat memudahkan untuk pencatatan transaksi, menganalisis pola pengeluaran, dan menyediakan gambaran menyeluruh kondisi keuangan[2]. Sehingga membantu membangun sistem terorganisir dalam mengatur pendapatan pengeluaran serta tabungan agar kondisi keuangan tetap sehat.

Aplikasi Money Lover merupakan sebuah aplikasi pengelolaan keuangan yang menyediakan fitur seperti pencatatan transaksi, pengingat tagihan, dan perencanaan anggaran yang dapat mengelola anggaran sehari – hari. Dengan adanya aplikasi ini, pengguna diharapkan mampu lebih bijak dalam mengelola keuangan secara praktis dan efisien. Namun efektivitas dan keberhasilan aplikasi tidak hanya bergantung pada fungsinya, tetapi juga ditentukan oleh tingkat kepuasan pengguna[3]. Salah satu indikator kepuasan pengguna dapat dilihat dari ulasan pada platform distribusi Google Play Store[4]. Ulasan ini berisi opini, komentar maupun kritik dari pengguna terkait pengalaman menggunakan aplikasi.

Jumlah ulasan yang sangat banyak, sifat ulasan berbentuk teks cenderung bersifat subjektif dan beragam membuat analisis manual menjadi sulit dilakukan. Sehingga diperlukan metode yang tepat untuk mengolahnya menjadi informasi yang lebih tersruktur[5]. Maka dari itu analisis sentimen memiliki peran penting, yaitu dengan teknik pemrosesan *Natural Language Procesing* (NLP) bertujuan untuk mengelompokkan dna menemukan opini-opini pengguna ke kategori sentimen tertentu, seperti positif dan negatif[6]. Dengan menggunakan metode analisis sentimen dapat secara mudah mengidentifikasi dan tren dalam ulasan pengguna yang berpotensi untuk memberikan informasi bagi pengembang untuk mengoptimalkan performa aplikasi[7]. Berbagai penelitian sebelumnya telah membahas analisis sentimen berbagai macam aplikasi, namun kajian yang secara khusus pada aplikasi pengelolaan keuangan terutama aplikasi Money Lover masih relatif terbatas, mengingat penting nya pengelolaan finansial, sehingga terdapat kesenjangan literatur yang ada.

Random Forest adalah algoritma klasifikasi yang bekerja melalui menggabungkan beberapa pohon keputusan supaya hasil akurasi prediksi model meningkat. *Random forest* dipilih me miliki keunggulan dalam mengatasi masalah *overfitting*[8]. Disisi lain, *Naive Bayes* meruapakan algorima klasifikasi yang bekerja dengan menghitung proprabilitas dan statistik berdasarkan teorema bayes. Algoritma *Naïve Bayes* dipilih karena pemrosesan yang cepat dan akurasi yang cukup tinggi[9].

Penggunaan kedua algortima ini dalam analisis sentimen pada aplikasi keuangan memungkinkan pemahaman yang lebih baik mengenai bagaimana pola opini pengguna dapat di klasifikasikan secara otomatis Pendekatan tersebut memungkinkan eksplorasi terhadap sejauh mana algoritma berbasis *ensemble* dan algoritma probabilistik memberikan hasil yang berbeda dalam memproses data ulasan yang bersifat subjektif dan beragam. Sekaligus dapat membuka peluang pada pengembangan analisis sentimen pada domain aplikasi finansial yang masih relatif terbatas.

2 Tinjauan Literatur

Analisis sentimen termasuk dalam bidang *Natural Language Procesing* (NLP) yang bertujuan untuk mengekstrak, mengolah, dan mengelompokkan opini dari ulasan pengguna aplikasi[10]. Dalam konteks Money Lover, analisis sentimen membantu memberi wawasan terkait pengalaman pengguna dan dapat digunakan sebagai landasan perbaikan sistem layanan atau pengembangan sebuah fitur. Beberapa penelitian terdahulu yang berkaitan dengan analisis sentimen dengan algoritma *Random Forest* dan *Naïve Bayes* telah dilakukan.

Penelitian yang dilakukan oleh Sianipar dkk, meneliti analisis sentimen aplikasi Dana dan LinkAja di Google Play Store dengan algoritma *Naive Bayes*. Hasil penelitian menunjukkan akurasi sebesar 90%, menegaskan efektivitas algoritma *Naive Bayes* dalam mengklasifikasikan ulasan pengguna. Namun, penelitian ini hanya melaporkan akurasi dan belum menampilkan metrik lain seperti presisi, *recall*, atau *F1-Score,* yang penting untuk mengevaluasi keseimbangan model pada data yang cenderung tidak seimbang[11].

Sementara itu, pada penelitian yang dilakukan Pambudi dkk, menganalisis sentimen pengguna aplikasi Fizzo Novel menggunakan kerangka SEMMA dengan algoritma *Support Vector Machine* (SMV) dan *Naive Bayes*. *Dataset* yang digunakan sebesar 139.759 ulasan dari Google Play Store. Hasil penelitian menujukkan bahwa SVM memiliki performa yang jauh lebih unggul dengan akurasi

 ISSN:2302-8149
e-ISSN:2540-9719

97,10% sebelum SMOTE dan 96,99% setelah SMOTE, serta F1-Score yang seimbang pada semua kelas sentimen. Sebaliknya, *Naive Bayes* hanya memperoleh akurasi 75,82% sebelum SMOTE dan menurun menjadi 73,63% setelah SMOTE. Temuan ini menegaskan keunggulan SVM dalam menangani data ulasan yang besar dan tidak seimbang[12].

Pada studi lainya oleh Kaeren dan Andrianingsih meneliti analisis sentimen aplikasi LinkAja di Google Play Store dengan membandingkan algoritma *Naive Bayes* dan *Random Forest*. *Dataset* yang digunakan mencakup lebih dari 1.800 ulasan. Hasil klasifikasi menunjukkan bahwa Random Forest unggul dengan akurasi 82%, presisi 84%, *recall* 81%, dan *F1-score* 82%, sedangkan *Naive Bayes* mencapai akurasi 79%. Penelitian ini juga menemukan bahwa *Random Forest* lebih efektif dalam mengidentifikasi sentimen negatif yang mendominasi ulasan aplikasi, sementara ulasan positif umumnya berhubungan dengan kemudahan penggunaan[13].

Penelitian oleh Larasati dkk, menggunakan *Random Forest* untuk analisis sentimen ulasan aplikasi Dana yang dikumpulkan melalui scraping di Google Play Store. Hasil penelitian menunjukkan bahwa dengan parameter terbaik (400 trees, depth 65), diperoleh akurasi 84%, presisi 84%, *recall* 84%, dan *F1-Score* 84%. Penelitian ini juga menyoroti dampak ketidakseimbangan kelas, di mana *recall* pada kelas negatif lebih rendah. Setelah dilakukan penyeimbangan data, performa model meningkat signifikan hingga mencapai akurasi 84% secara konsisten[14].

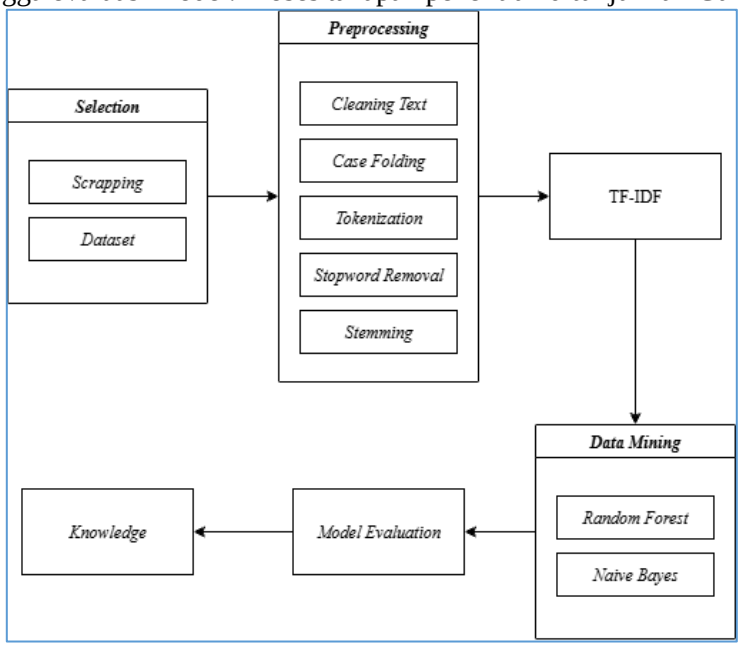
Dari penelitian terdahulu telah mengkaji metode – metode dalam analisis sentimen, studi yang membahas analisis sentimen ulasan pada aplikasi pengelolaan keuangan masih jarang ditemukan. Penelitian ini memberikan kontribusi pada analisis sentimen pada aplikasi finansial secara khusus aplikasi Money Lover, dan juga membandingkan dan menilai performa algoritma *Random Forest* dan *Naive Bayes* berdasarkan metrik evaluasi seperti akurasi, presisi, *recall*, *F1-score* dan *confusion matrix*. Perbandingan performa model machine learning ditunjukkan pada Tabel 1 berikut.

Tabel 1 Perbandingan penelitian terdahulu

Objek	Metode	Akurasi	Presisi	Recall	F1-Score
Dana dan LinkAja	Naive Bayes	90%	-	-	-
Fizzo Novel	Naive Bayes	74%	-	-	-
LinkAja	Random Forest	82%	84%	81%	82%
Dana	Random Forest	84%	84%	84%	84%

3 Metode Penelitian

Penelitain ini menerapkan metode yang meliputi tahapan – tahapan penting mulai dari pengumpulan data hingga evaluasi model. Proses tahapan penelitian ditunjukkan Gambar 1. berikut.



Gambar 1 Metode penelitian

 ISSN:2302-8149
e-ISSN:2540-9719

3.1 *Selection*

Tahap awal penelitian adalah memilih dan menyiapkan sumber data yang akan dianalisis. Data tersebut berupa ulasan atau komentar pandangan pengguna aplikasi mengenai kualitas layanan aplikasi Money Lover yang terdapat pada Google Play Store

3.1.1 *Scrapping*

Pada tahap pertama penelitian ini merencanakan pengumpulan data ulasan aplikais Money Lover yang didapatkan melalui Google Play Store dengan memanfaatkan teknik *web scrapping*. Data yang akan didapatkan mencakup teks ulasan yang memuat opini pengguna terkait aplikasi, serta rating (*score*) menggunakan pemrograman *Python*.

3.1.2 *Dataset*

Hasil dari proses scrapping dikumpulkan menjadi satu *dataset*. *Dataset* ini berisi kumpulan ulasan pengguna aplikasi Money Lover yang akan di gunakan pada proses *preprocessing* evaluasi model. Data yang talah diperoleh melalui Google Play Store meliputi 3000 dataset yang terdiri dari kolom *username*, *at*, *content*, *score*.

3.2 *Preprocessing*

Tahap ini bertujuan untuk mengolah data ulasan yang diperoleh agar teks ulasan berada dalam format bersih dan siap digunakan evaluasi model[15]. Tujuan tahapan ini adalah menghapus *noise*, menstandarisasi, dan menyiapkan teks agar model dapat mengenali pola sentimen secara optimal[16]. Data yang sebelumnya berjumlah 3000 data, setelah dilakukan *preprocessing data* menghasilkan 2877 data.

3.2.1 *Cleaning Text*

Pada tahap awal *preprocessing* adalah *data cleaning*. Teks ulasan akan dibersihkan dari informasi yang tidak relevan atau berpotensi menghambat proses analisis, meliputi angka, URL, tanda baca, emoji dan simbol khusus[17].

3.2.2 *Case Folding*

Dalam tahap ini seluruh teks ulasan akan di ubah ke huruf kecil untuk menyamakan kapitalisasi dari perbedaan yang disebabkan oleh variasi huruf besar. Tahap ini membantu konsistensi data sehingga model tidak salah menginterpretasi kata-kata yang identik[17].

3.2.3 *Tokenization*

Pada proses ini kalimat utuh akan dipecah menjadi satuan kata atau token. Tokenisasi memungkinkan algoritma untuk menganalisis kata per kata dan mempelajari hubungan antar kata dalam konteks sentimen[17].

3.2.4 *Stopword Removal*

Langkah ini mencakup penghapusan kata berdasarkan *stopword*, yaitu kata – kata pada dokumen yang bersifat umum dan tidak memiliki nilai penting untuk kontribusi analisis sentimen, seperti “untuk”, “di”, ”yang”, “dan” akan dihapus. Penghapusan *stopword* membantu mengurangi *noise* dalam data dan meningkatkan kualitas representasi teks.

3.2.5 *Stemming*

Proses terakhir *preprocessing* dilakukan *stemming* yaitu proses dimana setiap kata akan dikembalikan ke bentuk dasarnya. *Stemming* bertujuan agar kata-kata dengan makna sama diperlakukan seragam, sehingga meningkatkan konsistensi data dan efisiensi model[17].

3.3 TF – *IDF*

Pada penelitian ini, data ulasan akan mentraformasikan mejadi bentuk angka melalui penerapan teknik *Term Frequency-Inverse Document Frequency*(TF-IDF). Metode tersebut menetapkan nilai bobot pada setiap kata pada frekuensi kemunculanya di seluruh dokumen yang dianalisis [17]. Berikut rumus TF-IDF pada persamaan (1)

$$TF - IDF = idf_{td} \times idf_t \quad (1)$$

Di mana : *Term frequency* (TF) berfungsi menghitung intensitas kemunculan sebuah kata dalam suatu dokumen, dihitung pada persamaan (2) berikut.

$$TF = \frac{jumlah\ kemunculan\ kata\ dalam\ dokumen}{total\ kata\ dalam\ dokumen} \quad (2)$$

 ISSN:2302-8149
e-ISSN:2540-9719

Inverse Document Frequency (IDF) menghitung tingkat keunikan kata di keseluruhan korpus dihitung pada persamaan (3) berikut.

$$idf = \log\left(\frac{N}{dft}\right) \quad (3)$$

3.4 *Data Mining*

Tahapan *Data Mining* bertujuan untuk melatih dan menguji model *machine learning* melalui klasifikasi komentar pengguna, yang dibagi menejadi kategori positif dan negatif. Kumpulan data dibagi 80% digunakan sebagai pelatihan dan 20% sisanya untuk pengujian. *Random Forest* dan *Naive Bayes* adalah dua algoritma yang dibandingkan dalam penelitian ini.

3.4.1 *Random Forest*

Algoritma *Random Forest* adalah algoritma klasifikasi berbasis ensembel yang membagun pohon keputusan secara acak, kemudian menggabungkan prediksi mereka untuk menghasilkan prediksi yang lebih akurat dan tahan terhadap *overfit*. Masing-masing pohon dibuat berdasarkan data dan karateristik secara acak, dan keputusan akhir ditentukan oleh suara mayoritas dari semua pohon[18]. Pada algoritma *Random Forest* dikonfigurasi dengan menggunakan parameter *n_estimator=100* dan menetapkan random_state==42 guna menjaga konsistensi model. Perhitungan *Random Forest* pada persamaan (4) sebagai berikut:

$$H(x) = \arg max \sum_{i=1}^{k} I(h_i(x) = Y) \quad (4)$$

3.4.2 *Naive Bayes*

Algoritma *Naive Bayes* merupakan algoritma statistik sederhana namun sangat cepat dan efisien dalam mengklasifikasikan teks, khususnya pada *dataset* besar. Algoritma Naïve Bayes menawarkan karakteristik yang mirip dengan *decision tree*[19]. Pada algoritma *Naive Bayes* dikonfigurasikan dengan parameter defaultnya, dengan nilai *alpha=1.0* dan parameter *fit_prior=True* agar model secara otomatis menyesuaikan distribusi kelas. Algoritma *Naïve Bayes* berdasarkan pada teorama bayes dengan persamaan (5) sebagai berikut.

$$P\ (C|X) = \frac{P(X|C)}{P(X)} \quad (5)$$

3.5 Evaluasi Model

Tahap evaluasi dilakukan setelah model dilatih dan digunakan untuk memprediksi sentimen pada data *testing* yang sebelumnya tidak digunakan dalam pelatihan, sehingga hasil evaluasi mencerminkan kemampuan model dalam menghadapi data baru. Evaluasi didasarkan pada metrik akurasi, presisi, *recall* dan *f1-score*, untuk mengukur efektivitas dan kinerja model. Selain itu performa model juga divalidasi Berikut adalah metrik evaluasinya.

3.5.1 Akurasi

Metrik ini akan mengukur persentase seberapa besar hasil prediksi benar dari seluruh prediksi yang diklasifikasikan oleh model, dengan rumus persamaan (6) berikut.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

3.5.2 Presisi

Metrik ini akan mengukur tingkat model dalam melakukan prediksi yang tepat pada kedua kelas yaitu negatif dan positif, dengan rumus persamaan (7) berikut.

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

3.5.3 *Recall*

Metrik ini akan mengukur seberapa benar prediksi pada ketiga kelas yaitu negatif dan positif, dengan rumus persamaan (8) berikut.

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

3.5.4 *F1-Score*

Metrik ini merupakan hasil perhitungan nilai rata-rata dari *presisi* dan *recall*, dengan rumus persamaan (9) berikut.

$$\text{F1 } Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

 ISSN:2302-8149
e-ISSN:2540-9719

3.5.5 *Cross Validation*

Cross-validation merupakan teknik evaluasi yang melibatkan pembagian data menjadi beberapa bagian(*fold)*, kemudian melatih dan menguji model pada sebagian data dan diuji pada bagian lain secara bergantian, sehingga hasil evaluasi menjadi stabil dan representatif[20]. Pada penelitian ini, evaluasi model dilakukan dengan kumpulan data dibagi menjadi sepuluh bagian menggunakan prosedur k-fold cross validation dimana nilai k sama dengan 10, model akan dilatih pada 9 *fold* dan diuji pada 1 *fold* berbeda pada setiap iterasi yang bertujuan untuk memberikan estimasi kinerja model yang lebih akurat dan mengurangi bias.

3.5.6 *ROC Curve*

ROC Curve merupakan representasi grafis yang menunjukkan seberapa efektif model dapat membedakan kasus kelas positif dan negatif melalui evaluasi tingkat positif yang benar dan tingkat positif yang salah pada ambang tertentu. Nilai luas di bawah kurva (AUC) adalah ukuran yang menunjukkan seberapa baik model dalam klasifikasi secara keseluruhan[21]. dapat dihitung dengan persamaan 10 berikut.

$$AUC = \sum \left(\frac{TPR_i TPR_{i-1}}{2} \right) \times (FPR_i - FPR_{i-1}) \quad (10)$$

3.6 Pengetahuan

Tahap akhir penelitian menyimpulkan temuan dari hasil evaluasi model klasifikasi, termasuk model yang paling efektif mengenali sentimen Money Lover. Selain menilai performa model melalui metrik akurasi, presisi, *recall*, dan *F1-score*, juga memberikan wawasan mengenai pola sentimen pengguna, misalnya mayoritas ulasan positif atau negatif, serta tema atau kata yang sering muncul. Informasi ini dapat digunakan untuk memahami persepsi pengguna, memperbaiki kualitas aplikasi, dan menjadi dasar dalam pengambilan keputusan didasarkan pada analis data.

4 Hasil dan Pembahasan

Setiap tahapan dalam penelitian ini dijelaskan dengan logis dan teratur untuk mempermudah pemahaman yang jelas mengenai proses dan hasil yang dicapai, tahapan dimulai dari proses pengumpulan data, hingga penerapan model *data mining* dan evaluasi kinerja model. Pada bagian ini juga dipaparkan hasil yang diperoleh pada setiap tahapan, pembahasan terkait performa masing – masing metode berdasarkan metrik evaluasi.

4.1 Pengumpulan Data

Ulasan pengguna aplikasi Money Lover di Google Play Store merupakan data yang akan digunakan pada penelitian ini. Untuk memudahkan pengumpulan data, digunakan metode web scraping dan bahasa *Python* dengan pustaka google-play-scapper.

userName	at	content	score	sentimen
Naufal Alif Ganistra	2025-11-06 06:43:24	saya pengguna money lover sudah 5 tahun namun baru kali ini data keuangan ketika saya buka di android dan iphone nilainya berbeda, pdhl dengan akun yang sama apakah ada kesalahan sistem atau memng seperti itu jika menggunakan 2 gawai berbeda	4	positif
Sultan DA	2025-10-26 16:27:08	ini kalau beli di android, di ios apa otomatis premium juga?	4	positif
Febriani Feby	2025-10-22 02:58:05	sangat membantu dalam mengelolah keuangan	5	positif
Jefri Dwi	2025-10-15 12:49:58	bagus, dan mudah aplikasinya	5	positif
Juanda Mahendra	2025-10-13 06:16:21	Simple, Mudah Digunakan, Siapapun Orang Awam Mudah Mengerti Akan Apk Ini.	5	positif
...	...	...	...	...
Pengguna Google	2019-04-16 02:21:09	semoga gak ada iklan yang merusak aplikasi yg sudah bagus ini	5	positif
Pengguna Google	2019-04-15 13:02:52	membantu sekali untuk mengelola keuangan saya sehari-hari	5	positif
Pengguna Google	2019-04-15 09:49:01	bagus& sangat membantu	5	positif
Pengguna Google	2019-04-15 01:43:28	jadi rajin menabung saya	5	positif
Pengguna Google	2019-04-14 16:56:06	sangat membantu buat mengatur keuangan (apalagi anak kos), gampang pake ga ribet, fiturnya lengkap	5	positif

Gambar 2 Data ulasan hasil scraping data

Berdasarkan Gambar 2 tersebut proses *scraping* data yang dilakukan mencakup kolom username, at, content, score, dengan jumlah data awal sebesar 3000 ulasan, kemudian akan diberi 2 kategori label sentimen yaitu positif dan negatif. Kategori sentimen ditentukan dengan kolom score 1 dan 2 sebagai negatif, score 4 dan 5 sebagai positif.

4.2 *Preprocessing* Data

Tahap *preprocessing* data merupakan tahap penting untuk memastikan kualitas dan kebersihan data. Tujuannya adalah menghilangkan elemen – elemen tidak relevan, *noise* dan menyiapkan data

 ISSN:2302-8149
e-ISSN:2540-9719

teks menjadi lebih terstruktur. Langkah pertama adalah pembersihan teks (*cleaning text*). Selama proses ini, seluruh karakter non-tekstual antara lain tanda baca, emoji, angka, dan simbol dihapus.

Tahap berikutnya adalah *case folding*, proses ini memodifikasi format keseluruhan teks ke huruf kecil dan mencegah model memperlakukan kata yang sama dengan kapitalisasi berbeda sebagai dua entitas terpisah. Setelah normalisasi awal tahap selanjutnya dilakukan *tokenization*, yaitu tahapan untuk membagi teks menjadi potongan-potongan kata. Setiap potogan kata kemudian melewati tahap *filtering* dengan menggunakan *stopword remova*l. Kata yang dihapus adalah kata-kata yang bersifat umum dan tidak memiliki kontribusi makna untuk analisis sentimen.

Setelah proses *tokenization* dan *stopword removal*, proses berikutnya adalah *stemming*. Dengan menggunakan pustaka sastrawi, proses ini bertujuan mengembalikan kata-kata ke bentuk dasarnya. Data yang sebelumnya berjumlah 3000 data, setelah dilakukan *preprocessing data* menghasilkan 2877 data. Hasil dari *preprocessing* ditunjukkan pada Gambar 3 berikut

content	cleaning_text	case_folding	tokenization	stopword_removal	stemming
saya pengguna money lover sudah 5 tahun namun ...	saya pengguna money lover sudah tahun namun ba...	saya pengguna money lover sudah tahun namun ba...	[saya, pengguna, money, lover, sudah, tahun, n...	[pengguna, money, lover, tahun, baru, kali, da...	[guna, money, lover, tahun, baru, kali, data, ...
ini kalau beli di android, di ios apa otomatis...	ini kalau beli di android di ios apa otomatis ...	ini kalau beli di android di ios apa otomatis ...	[ini, kalau, beli, di, android, di, ios, apa, ...	[kalau, beli, android, ios, apa, otomatis, pre...	[kalau, beli, android, ios, apa, otomatis, pre...
sangat membantu dalam mengelolah keuangan	sangat membantu dalam mengelolah keuangan	sangat membantu dalam mengelolah keuangan	[sangat, membantu, dalam, mengelolah, keuangan]	[sangat, membantu, mengelolah, keuangan]	[sangat, bantu, elo, uang]
bagus, dan mudah aplikasinya	bagus dan mudah aplikasinya	bagus dan mudah aplikasinya	[bagus, dan, mudah, aplikasinya]	[bagus, mudah, aplikasinya]	[bagus, mudah, aplikasi]
Simple, Mudah Digunakan, Siapapun Orang Awam M...	simple mudah digunakan siapapun orang awam mud...	simple mudah digunakan siapapun orang awam mud...	[simple, mudah, digunakan, siapapun, orang, aw...	[simple, mudah, digunakan, siapapun, orang, aw...	[simple, mudah, guna, siapa, orang, awam, muda...

Gambar 3 Hasil *preprocessing data*

4.3 Pembobotan TF-IDF

Setelah tahap preprocessing, data tersebut dikonversi menjadi bentuk representasi vektor numerik dengan menerapkan teknik *Term Frequency-Inverse Document Frequency*(TF-IDF). Selanjutnya hasil dari trasnformasi TF-IDF ditunjukkan pada Gambar 4 berikut

term	TF	IDF	TF-IDF
bagus	157.78108324122937	2.891073372185102	157.78108324122937
bantu	144.0859825186353	2.84153923230853	144.0859825186353
sangat	140.37599383987904	2.81229607800689	140.37599383987904
good	140.26112882840546	5.572469536142939	140.26112882840546
sangat bantu	113.39398749123313	3.05791256476409	113.39398749123313
aplikasi	81.53020661088442	2.704094825446417	81.53020661088442
uang	71.81580719377151	2.3025841188940032	71.81580719377151
mantap	69.75778240210742	4.21105773245244?	69.75778240210742
keren	42.37635373675685	4.507626752321475	42.37635373675685
banget	39.99235829258163	3.90149294875119?	39.99235829258163

Gambar 4 Hasil dari TF-IDF

Pada hasil penerapan TF-IDF, teks yang telah diolah menjadi representasi vektor numerik dengan menggunakan *Tfidfvectorizer* dari pustaka *scikit-learn*. Hasil akhir dari proses ini adalah sebuah matriks yang akan digunakan dalam analisis sentimen menggunakan algoritma *data mining*.

4.4 *Data Mining*

Setelah data melalui tahap *preprocessing*, data ulasan aplikasi Money Lover yang telah bersih dan diberi label. Data yang berjumlah 2.877 data akan dibagi 80% sebagai pelatihan dan sisanya 20% sebagai pengujian, pembagian menghasilkan 2.302 data pelatihan dan 575 data pengujain. Selanjutnya data akan diproses yang melibatkan implementasi *Random Forest* dan *Naive Bayes*, menggunakan fitur yang telah dibobotkan oleh TF-IDF. Pada tahap ini, model dilatih untuk mengidentifikasi pola pembelajaran yang dapat diguanakan untuk mengklasifikasikan data uji berdasarkan sentimen positif atau negatif.

4.4.1 *Random Forest*

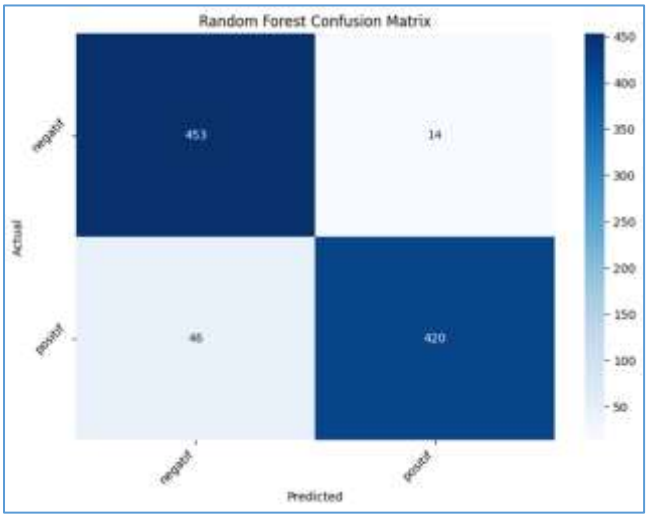
Setelah model dilatih dengan *dataset*, evaluasi model dilakukan untuk mengukur kemampuan model dalam mengklasifikasikan sentimen ulasan aplikasi Money Lover. Hasil pengujian *Random Forest* ditampilkan pada Gambar 5. berikut.

 ISSN:2302-8149
e-ISSN:2540-9719

	precision	recall	f1-score	support
negatif	0.91	0.97	0.94	467
positif	0.97	0.90	0.93	466
accuracy			0.94	933
macro avg	0.94	0.94	0.94	933
weighted avg	0.94	0.94	0.94	933

Gambar 5 Metrik *random forest*

Berdasarkan Gambar tersebut hasil pengujian model, nilai presisi pada kelas negatif mencapai 91% yang berarti sebagian besar prediksi negatif adalah benar, untuk *recall* negatif memiliki nilai 97% menandakan kemampuan tinggi model dalam menangkap data negatif secara benar. Sementara kelas positif memiliki presisi 97% yang menunjukkan tingkat kesalahan prediksi sangat rendah, pada recall kelas positif memiliki nilai 90%. Nilai *f1-score* pada kelas positif 93% dan pada kelas negatif 94%, menandakan keseimbangan presisi dan *recall*. Secara keseluruhan model *Random Forest* menunjukkan nilai akurasi mencapai 94% menunjukkan bahwa total prediksi model yang sudah benar. Untuk memberikan lebih detail gambaran prediksi ditunjukkan *Confusion Matrix Random Forest* pada Gambar 6 berikut.



Gambar 6 Confusion matrix *random forest*

Berdasarkan *confusion matrix* model *Random Forest*, dari data ulasan negatif diklasifikasikan 453 data benar negatif, hanya 14 yang salah di prediksi sebagai positif, menandakan kemampuan model yang sangat baik dalam mengklasifikasikan ulasan negatif. Untuk ulasan positif diklasifikasikan 420 data benar positif, sementara 46 data salah teriidentifikasi sebagai negatif. Hal ini menandakan kemampuan model sangat baik dalam mengklasifikasikan sentimen.

4.4.2 *Naïve Bayes*

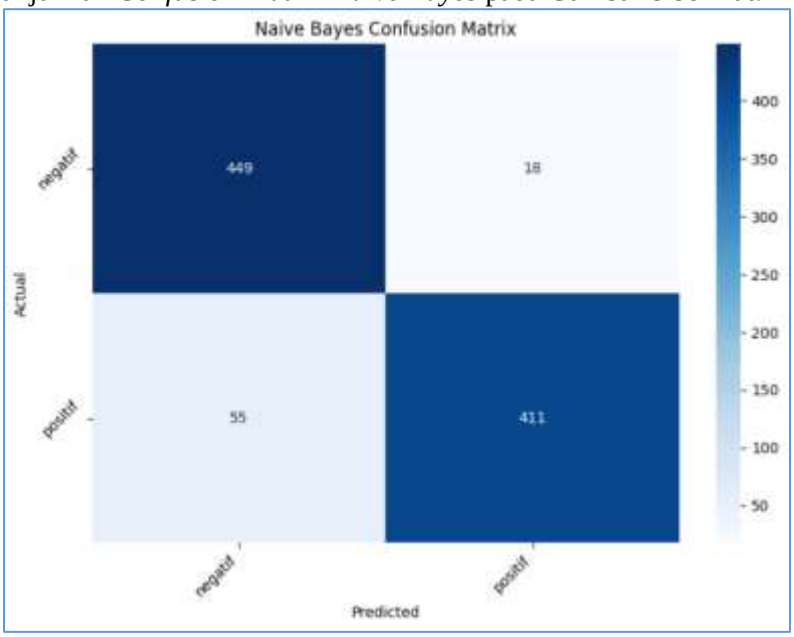
Pada pengujian model *naive bayes* dengan data ulasan Money Lover menunjukkan hasil evaluasi model dengan nilai akurasi mencapai 92%. Ditunjukkan pada Gambar 7 berikut.

	precision	recall	f1-score	support
negatif	0.89	0.96	0.92	467
positif	0.96	0.88	0.92	466
accuracy			0.92	933
macro avg	0.92	0.92	0.92	933
weighted avg	0.92	0.92	0.92	933

Gambar 7 Metrik naive bayes

 ISSN:2302-8149
e-ISSN:2540-9719

Berdasarkan Gambar tersebut, model *Naive Bayes* menunjukkan nilai presisi pada kelas negatif 89%, pada recall negatif memiliki nilai 96% yang berarti sebagian besar data negatif berhasil diidentifikasi dengan benar oleh model. Pada kelas positif nilai presisi mencapai 96% namun nilai recall lebih rendah sebesar 88% menunjukkan bahwa masih terdapat beberapa data positif yang salah terklasifikasi. Nilai f1-score untuk kedua kelas berada pada angka 0.92, menandakan keseimbangan performa model dalam memprediksi kedua kategori sentimen. Untuk memberikan gambaran lebih detail prediksi ditunjukkan *Confusion Matrix Naive Bayes* pada Gambar 8 berikut.



Gambar 8. Confusion matrix naive bayes

Berdasarkan *confusion matrix Naive Bayes*, data ulasan negatif model mengklasifikasikan 449 secara benar sebagai negatif, hanya 18 yang salah diprediksi sebagai positif, menandakan kemampuan model cukup baik untuk mengenali data ulasan negatif. Pada ulasan positif sebanyak 411 data diprediksi benar sebagai positif, namun terdapat 55 data lainnya yang salah diklasifikasikan sebagai negatif, namun terdapat 55 data salah diprediksi sebagai negatif. Secara keseluruhan pola prediksi yang ditampilkan masih konsisten dengan metrik evaluasi sebelumnya dan menunjukkan bahwa model bekerja cukup efektif dalam memproses ulasan aplikasi.

4.5 Evaluasi Model

Model akan dievaluasi dengan cara membandingkan metrik akurasi, preisis, recall, dan f1-score hasil dari kinerja model *Random Forest* dan *Naive Bayes*. Hasil evaluasi tertera pada Tabel 2 berikut.

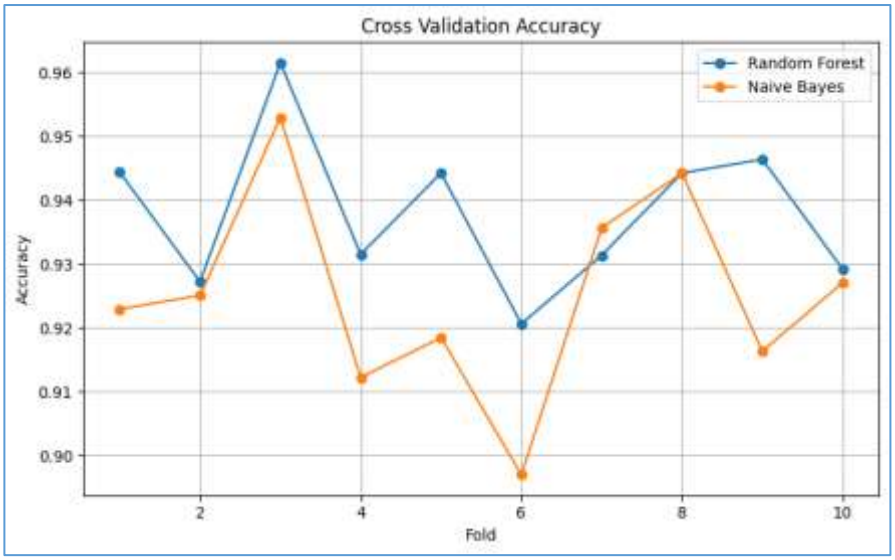
Tabel 2. Perbandingan metrik

Metrik	Random Forest	Naive Bayes
Akurasi	94%	92%
Presisi	94%	92%
Recall	94%	92%
F1-Score	94%	92%

Berdasarkan Tabel tersebut hasil pengujian model *Random Forest* menunjukkan performa angka presisi, *recall*, dan f1-score sebesar 94%, menandakan keseimbangan performa yang konsisten pada seluruh kelas. Sementara itu, model *Naive Bayes* menunjukkan nilai preisisi, *recall*, dan f1-score sebesar 92%, yang meskipun stabil, tetap lebih rendah dibandingkan model *Random Forest*.

Untuk melihat perbedaan performa model secara menyeluruh dan memperkuat temuan tersebut, dilakukan analisis tambahan menggunakan grafik *cross-validation* dan grafik *ROC Curve* sebagai alat evaluasi tambahan. Hasil dari *cross-validation* pada Gambar 9 berikut.

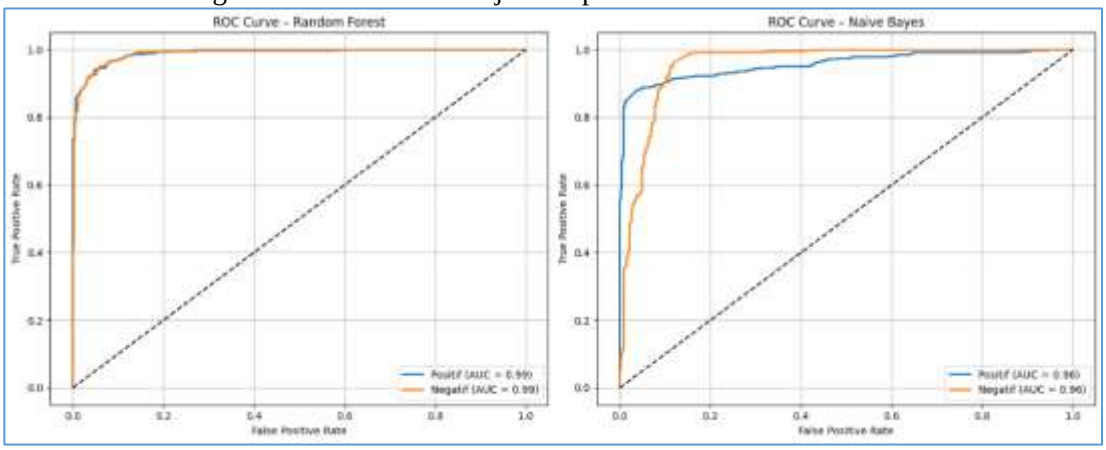
 ISSN:2302-8149
e-ISSN:2540-9719



Gambar 9 Grafik *cross-validation*

Berdasarkan Gambar 10 tersebut hasil *cross-validation*, menunjukkan model *Random Forest* memiliki nilai akurasi yang sebagian besar di atas 0.92 dengan nilai akurasi tertinggi pada 0.96 pada *Fold 3* dengan nilai fluktuasi akurasi yang tidak terlalu tajam. Hal ini menunjukkan kemampuan stabilitas model yang baik dan kinerja tinggi pada *subset* data. Sementara pada model Naive Bayes menunjukkan rentang nilai akurasi yang lebih beragam dengan nilai maksimum 0.95 dan nilai minimum 0.89, meskipun kemampuan model cukup baik dan mampu mencapai performa tinggi pada *Fold 3* dan *Fold 8*, namun nilai akurasinya lebih fluktuatif dan turun signifikan, menunjukkan model lebih sensitif terhadap perubahan distribusi data pada setiap *Fold*. Hasil *cross-validation* model *Random Forest* lebih konsisten dan unggul dalam performa rata – rata, sedangkan model Naive Bayes menunjukkan lebih fluktuatif meskipun dapat memberikan kinerja yang kompetitif pada kondisi tertentu

Setelah evaluasi performa melalui *cross-validation* untuk memastikan konsistensi hasil, dibuat grafik *ROC* Curve. Hasil grafik *ROC Curve* ditunjukkan pada Gambar 10 berikut.



Gambar 10 Grafik ROC *curve random forest, naive bayes*

Berdasarkan hasil visualisasi *ROC Curve* tersebut, model Random Forest menunjukkan nilai AUC 0.99 untuk setiap kelas yang menandakan kemampuan model dalam klasifikasi yang sangat baik dan hampir sempurna. Dengan demikian, model Random Forest bekerja dengan baik dalam hal pembagian kelas positif dan negatif selama seluruh fase klasifikasi. Model *Naive Bayes* menujukkan hasil yang cukup baik dengan AUC 0.96 untuk membedakan kategori positif dan negatif, meskipun kurvanya sedikit lebih landai dibanding *Random Forest*. Secara keseluruhan, kedua model memiliki kemampuan diskriminatif yang tinggi, namun model Random Forest lebih unggul karena menghasilkan area kurva yang lebih konsisten pada kedua kelas sentimen.

 ISSN:2302-8149
e-ISSN:2540-9719

4.6 Pengetahuan

Berdasarkan hasil evaluasi, kedua model tersebut berkinerja cukup baik dalam mengklasifikasi sentimen ulasan aplikasi Money Lover. Model *Random Forest* menunjukkan performa yang sedikit lebih unggul dibandingkan dengan *Naive Bayes*. Keunggulan ini karena model *Random Forest* merupakan metode berbasis *ensemble* berbasis pembentukan banyak pohon keputusan sehingga mampu lebih baik dalam menangkap variasi fitur teks, interaksi antar fitur serta ketidakteraturan data. pendekatan ini membuat model *Random Forest* lebih tahan terhadap noise dan mengurangi risiko overfitting, sehingga hasil klasifikasi menjadi konsisten pada setiap *fold cross-validation*. Pada model *naive bayes* bekerja dengan mengasumsikan kemandirian fitur, yang dalam konteks data teks sering kali tidak terpenuhi. Kata – kata dalam kalimat saling terkait secara semantik, sehingga pelanggaran asumsi ini dapat menyebabkan penurunan pada akurasi. Meskipun demikian kecepatan dan efisiensi lebih unggul dalam proses pelatihan, namun performa masih di bawah *Random Forest* dari hasil ROC maupun metrik evaluasi.

Analisis lebih lanjut mengungkapkan bahwa sebagian besar kesalahan klasifikasi pada kedua model terjadi pada ulasan yang mengandung sentimen ambigu, menggunakan kata kiasan, serta cenderung netral yang sulit dipetakan positif atau negatif, yang sulit diinterpretasikan bahkan oleh model canggih. Penelitian ini juga menghasilkan pemahaman yang lebih mendalam mengenai efektivitas algortima *machine learning* dalam tugas analisis sentimen dan mengisi celah penelitian yang ada.

5 Kesimpulan

Tujuan penelitian ini adalah membandingkan performa algoritma *Random Forest* dan *Naive Bayes* dalam menganalisis ulasan sentimen aplikasi Money Lover di Google Play Store. Sebelum proses klasifikasi, data diproses melalui tahap *preprocessing*, pembobotan TF-IDF, dan penerapan kedua algoritma tersebut. Hasil evaluasi dapat dilihat *Random Forest* memiliki performa yang lebih baik dengan nilai akurasi 94% sedangkan *Naive Bayes* hanya mendapatkan nilai akurasi sebesar 92%. Perbedaan hasil ini mengindikasikan bahwa pendekatan dengan metode *ensemble* memiliki performa lebih unggul dan optimal dalam tugas klasifikasi pada data ulasan Money Lover. Temuan ini juga diperkuat dengan hasil dari *ROC Curve* dan *cross-validation* yang menunjukkan kestabilan serta kinerja *Random Forest* jauh lebih baik dari *Naive Bayes*. Penelitian ini juga dapat digunakan sebagai rujukan terkait efektivitas kedua algoritma dalam memproses ulasan pengguna aplikasi pengelolaan keuangan. Selain itu hasil penelitian ini membuka ruang untuk penelitian selanjutnya mengenai pola sentimen pada aplikasi finansial, mengingat pentingnya pengelolaan keuangan dalam kehidupan sehari-hari. Pengembangan lebih lanjut masih memungkinkan, seperti menggunakan jumlah data yang lebih besar dan variatif, eksplorasi teknik *feature selection* lainnya dan evaluasi model berbasis *cost-sensitive* analisis.

Referensi

[1] J. H. Napitupulu, N. Ellyawati, dan R. F. Astuti, “Pengaruh Literasi Keuangan dan Sikap Keuangan Terhadap Perilaku Pengelolaan Keuangan Mahasiswa Kota Samarinda,” Vol. 9, No. 3, 2021.

[2] I. Jayanto, A. Lubis, R. Hamzah, H. Indah, dan R. Ningtyas, “Penerapan Aplikasi Pencatatan Keuangan Digital bagi Ibu Rumah Tangga di Perumahan Mekarsari , Kota Depok (Literasi Keuangan menggunakan Aplikasi Money Manager dan Excel Sederhana),” Vol. 3, No. 4, hal. 277–285, 2024.

[3] M. Arsadhana, B. Efendi, dan M. Trihudiyatmanto, “Analisis Kepuasan Pelanggan Melalui Sentimen Ulasan menggunakan *Algoritma Naive Bayes*,” Vol. XIII, No. 1, hal. 1–8, 2025.

[4] R. Maulana, A. Voutama, dan T. Ridwan, “Analisis Sentimen Ulasan Aplikasi Mypertamina pada *Google Play Store* menggunakan Algoritma NBC,” *J. Teknol. Terpadu*, Vol. 9, No. 1, hal. 42–48, 2023.

[5] A. Sitanggang, Y. Umaidah, R. I. Adam, U. S. Karawang, dan T. Timur, “Analisis Sentimen Masyarakat Terhadap Program Makan Siang Gratis pada Media Sosial X menggunakan *Algoritma Naïve Bayes*,” Vol. 12, No. 3, 2024.

[6] R. C. Rivaldi dan T. D. Wismarini, “Analisis Sentimen pada Ulasan Produk dengan Metode

 ISSN:2302-8149
e-ISSN:2540-9719

Natural Language Processing (NLP) (Studi Kasus Zalika Store 88 Shopee),” Vol. 17, No. 1, hal. 120–128, 2024.

[7] N. Aditiya, P. Setiaji, dan Supriyono, “Analisis Sentimen Kepuasan Masyarakat terhadap Aplikasi ‘ INFO BMKG ’ menggunakan *Naive Bayes , SVM ,* dan *KNN*,” Vol. 14, hal. 1418–1432, 2025.

[8] N. D. Kurniawan, P. R. Ferdian, dan N. Hidayati, “Analisis Sentimen *Algoritma Naïve Bayes , Support Vector Machine ,* dan *Random Forest* pada Ulasan Aplikasi Ajaib,” Vol. 01, hal. 87–97, 2025.

[9] N. Widiastuti, A. Hermawan, dan D. Avianto, “Implementasi Metode *Naïve Bayes* untuk Klasifikasi Data Blogger,” Vol. 8, No. 3, hal. 985–994, 2023.

[10] A. R. Ismail, “Analisis Sentimen Komentar Pengguna Aplikasi Akademik menggunakan *Naïve Bayes* dan Seleksi Fitur *Particle Swarm Optimization Sentiment Analysis of Academic Application User Comments using Naïve Bayes and Particle Swarm Optimization for Feature Selection*,” Vol. 14, hal. 2565–2574, 2025.

[11] A. N. Sianipar, Yuhelmi, dan M. Devega, “Analisis Sentimen Ulasan Pengguna Aplikasi Dana dan Linkaja menggunakan Metode *Naïve Bayes*,” Vol. 6, No. 2n, hal. 510–522, 2024.

[12] S. Pambudi, P. Setiaji, dan W. A. Triyanto, “*Sentiment Analysis of Fizzo Novel Application using Support Vector Machine and Naïve Bayes Algorithm with SEMMA Framework*,” Vol. 6, No. 4, hal. 1861–1880, 2025.

[13] Kaeren dan Andrianingsih, “Analisis Sentimen Aplikasi Linkaja di *Google Play Store* menggunakan *Algoritma Naïve Bayes* dan *Random Forest*,” Vol. 06, No. 02, hal. 438–447, 2025.

[14] F. A. Larasati, D. E. Ratnawati, dan B. T. Hanggara, “Analisis Sentimen Ulasan Aplikasi Dana dengan *Metode Random Forest*,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, Vol. 6, No. 9, hal. 4305–4313, 2022.

[15] R. Amelia dan E. Indra, “Analisis Sentimen Ulasan Produk menggunakan *Large Languange Models* : Studi Kasus pada Shopee,” *JUSIKOMP*, Vol. 8, No. 2, hal. 1–13, 2025.

[16] E. Septiana dan C. Caroline, “Optimalisasi Evaluasi Pelaksanaan Pelatihan melalui Analisis Sentimen Otomatis dengan Model *Text Classification,*” hal. 141–154, 2024.

[17] L. Agustin, M. Fahmi, R. Dahlia, dan M. Rifqi, “Analisis Ulasan Konsumen sebagai Data Non-Keuangan dalam Sistem Informasi Akuntansi,” Vol. 5, No. 1, hal. 64–74, 2025.

[18] T. P. Subandono, D. Ariatmanto, M. T. Informatika, F. I. Komputer, K. Sleman, dan D. Istimewa, “Optimalisasi Seleksi Fitur dalam Analisis Sentimen Bank Saqu : Studi Perbandingan *SVM* dan *Random Forest* menggunakan *Information Gain* dan *Chi - Square Optimizing Feature Selection in Sentiment Analysis of Bank Saqu : A Comparative Study of SVM and Random Forest using Information*,” Vol. 14, hal. 1205–1219, 2025.

[19] M. F. Alamsyah dan A. Wijaya, “Perbandingan Metode *KNN* dan *Naïve Bayes* dalam Deteksi Tingkat Stres berdasarkan Ekspresi Wajah,” Vol. 10, No. 2, hal. 359–369, 2025, DOI: 10.30591/jpit.v10i2.8513.

[20] A. I. Pradana dan V. Atina, “Teknik *K-Fold Cross Validation* untuk mengevaluasi Kinerja Mahasiswa,” hal. 239–248, 2024, DOI: 10.33364/algoritma/v.21-1.1618.

[21] B. A. Sadewa dan Y. Yamasari, “Implementasi *Deep Transfer Learning* untuk Klasifikasi Nominal Uang Kertas Rupiah,” Vol. 05, hal. 543–551, 2024.