

Analisis Sentimen Pengguna Terhadap Faktor Penggunaan PayPal Menggunakan Metode *Decision Tree*

Maharani¹, Fathoni²

^{1,2}Universitas Sriwijaya, Indonesia

¹09031381924100@student.unsri.ac.id, ²fathoni@unsri.ac.id

ABSTRAK. PayPal, dompet digital dengan 429 juta pengguna global, tidak termasuk dompet digital yang umum digunakan di Indonesia sehingga perlu analisis sentimen masyarakat Indonesia terhadap layanan PayPal. Tujuan penelitian ini adalah untuk mengetahui sentimen masyarakat Indonesia terhadap PayPal berdasarkan teori bauran pemasaran 4P dan dapat menjadi referensi untuk peningkatan daya saing di Indonesia. Penelitian ini menggunakan model *Decision Tree* dengan kriteria *gain ratio*, *information gain*, dan *gini index* serta pengujian *K-Fold Cross Validation* menggunakan nilai *fold*=2, 3, 4, dan 5 untuk memproses 20.000 data ulasan Google Play Store. Hasilnya menunjukkan akurasi tertinggi pada pengujian *5-Fold Cross Validation* dengan kriteria *gain ratio* memprediksi sentimen negatif terhadap aspek produk (97,33%), harga (99,68%), dan tempat (96,72%). Hal ini memberi PayPal peluang untuk memperbaiki kualitas ketiga aspek ini. Sementara itu, prediksi sentimen positif terhadap aspek promosi sebesar 62,18%, menandakan bahwa layanan promosi PayPal telah dinilai cukup baik.

Kata Kunci: PayPal; Analisis Sentimen; Teori Bauran Pemasaran 4P; *Decision Tree*; *K-Fold Cross Validation*

ABSTRACT. PayPal, a digital wallet with 429 million global users, is not widely utilized in Indonesia. Therefore, there is a need for sentiment analysis on how the Indonesian community perceives PayPal's services. This research aims to understand Indonesian sentiment toward PayPal using the 4P marketing mix theory, serving as a reference for enhancing competitiveness. This research employs a *Decision Tree* model with *gain ratio*, *information gain*, and *Gini index*, along with *K-Fold Cross Validation* (*fold*=2, 3, 4, 5) on 20,000 Google Play Store reviews. The results reveal the highest accuracy in the *5-Fold Cross Validation* with *gain ratio* criteria predicting negative sentiments towards product (97.33%), price (99.68%), and place (96.72%) aspect. This provides an opportunity for PayPal to improve these aspects and enhance user satisfaction. Meanwhile, the positive sentiment prediction for promotion aspect (62.18%) indicates satisfactory performance in PayPal's promotional services.

Keywords: PayPal; Sentiment Analysis; 4P Marketing Mix Theory; *Decision Tree*; *K-Fold Cross Validation*

PENDAHULUAN

PayPal adalah aplikasi dompet digital yang dikembangkan dari perusahaan PayPal di Amerika Serikat yang menyediakan layanan transaksi keuangan secara online (Hapsari et al., 2021). PayPal dapat digunakan antar berbagai negara dan memiliki jumlah pengguna sekitar 429 juta orang pada Januari 2022 sampai Maret 2022 (Widi, 2022). Meskipun demikian, PayPal belum termasuk ke dalam kelompok dompet digital yang umum digunakan di Indonesia berdasarkan survei yang dilakukan Jakpat dan Insight Asia (Prasasti, 2022; Respati, 2022).

PayPal perlu merencanakan strategi agar dapat bersaing di pasar Indonesia. Salah satu cara untuk membangun strategi pemasaran adalah dengan mengetahui bagaimana pandangan masyarakat Indonesia terhadap layanan yang ditawarkan PayPal melalui analisis sentimen. Analisis sentimen merupakan sebuah teknik percabangan dari bidang *text mining* yang secara otomatis mengerti, mengekstrak, dan memproses informasi tekstual untuk mengidentifikasi apakah sentimen tersebut positif, negatif, atau netral (Hoang et al., 2019; Sari & Wibowo, 2019). Tingkatan umum analisis sentimen adalah tingkatan dokumen, tingkatan kalimat, dan tingkatan aspek. Analisis sentimen berbasis aspek atau *Aspect Based Sentiment Analysis* (ABSA) merupakan suatu teknik yang memproses informasi rincian teks yang berhubungan dengan aspek tertentu yang diteliti (Hoang et al., 2019; Pamungkas & Romadhony, 2021).

Salah satu pendekatan yang umum digunakan pada analisis sentimen adalah pendekatan berbasis mesin belajar (*machine learning*) karena tingkat hasil proses memiliki keakuratan yang baik (Birjali et al., 2021). Salah satu pendekatan berbasis *machine learning* yang sering digunakan adalah *supervised learning*, pendekatan ini membutuhkan data hasil pelatihan yang telah memiliki label berdasarkan polaritasnya. Salah satu metode berbasis *supervised learning* adalah *Decision Tree*. *Decision Tree* digunakan pada penelitian ini karena metode ini mampu mengelola *dataset* yang cukup besar, memiliki fleksibilitas yang bagus, dan memiliki akurasi yang cukup baik (Birjali et al., 2021; Permana et al., 2021; Puspita & Widodo, 2021).

Pengujian model *Decision Tree* dilakukan menggunakan metode *K-Fold Cross Validation*. *K-Fold Cross Validation* merupakan metode evaluasi model yang umum digunakan. *Dataset* biasanya dibagi menjadi data latih dan data uji, dan metode *K-Fold Cross Validation* melibatkan pelatihan data latih sebanyak *k-fold* iterasi. Kelebihan dari *K-Fold Cross Validation* adalah penentuan model yang lebih akurat karena metode ini membagi data menjadi beberapa *fold* pengujian (Fuadah et al., 2022; Peryanto et al., 2020).

Aspek yang diteliti dalam penelitian ini dibagi berdasarkan teori bauran pemasaran 4P. Teori bauran pemasaran 4P adalah teori yang dikemukakan oleh McCarty pada tahun 1964 yang mengamati aspek dasar pemasaran yaitu *product* (produk), *price* (harga), *place* (tempat), dan *promotion* (promosi), agar mampu mewujudkan sasaran pemasaran (Daryanto et al., 2019; Lestari et al., 2019; Majidah & Istianah, 2023). Produk mengacu kepada segala bentuk benda yang dijual oleh perusahaan ke konsumen sehingga dapat memenuhi kebutuhan konsumen dan mencapai tujuan pemasaran. Harga mengacu kepada nilai tukar produk yang disesuaikan berdasarkan sasaran, permintaan, dan biaya produk sehingga mampu menghasilkan keuntungan yang maksimal. Tempat mengacu kepada lokasi pemasaran yang mencakup lokasi perusahaan, lokasi persediaan, dan saluran distribusi barang dengan tujuan memudahkan proses distribusi produk sehingga konsumen mampu dengan mudah melakukan transaksi pembelian. Promosi mengacu kepada kegiatan penjualan produk yang mencakup perencanaan, penerapan, dan pengendalian komunikasi perusahaan kepada konsumen agar pelanggan tertarik dengan produk yang ditawarkan sehingga mencapai peningkatan penjualan produk (Abdullah, 2019; Majidah & Istianah, 2023; Sepfiani & Nawawi, 2023; Zahrah et al., 2021).

Salah satu tempat untuk mengetahui sentimen atau pandangan masyarakat terhadap aplikasi PayPal adalah melalui Google Play Store. Google Play Store merupakan pusat distribusi resmi untuk pengguna perangkat Android yang menyediakan layanan terkait instalasi aplikasi mobile. Google Play Store juga menyediakan fitur ulasan dan peringkat suatu aplikasi. Pengguna dapat menuliskan pengalaman menggunakan aplikasi yang diunduh dan melihat ulasan oleh pengguna lain sehingga pengguna dapat memiliki informasi tentang suatu aplikasi yang akan diunduh (Arbaini, 2020; Saputra et al., 2019; Wisudawati et al., 2021). Pada Google Play Store, aplikasi PayPal diunduh lebih dari 100 juta kali dan jumlah ulasan aplikasi PayPal pada *smartphone* mencapai 2,91 juta ulasan.

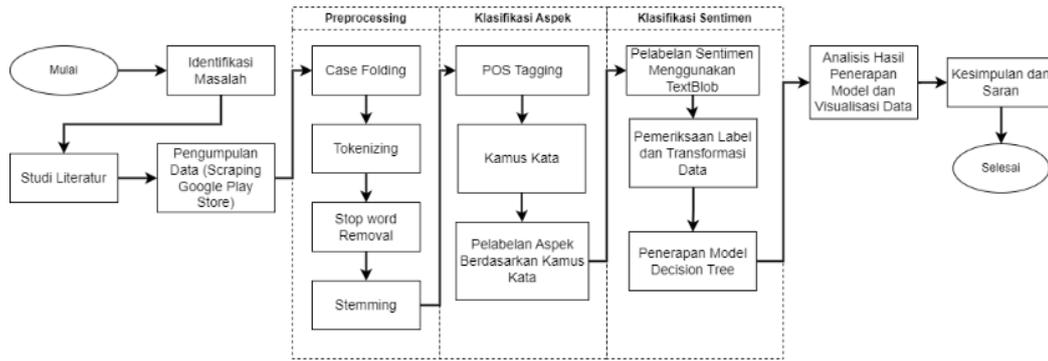
Penelitian tentang analisis sentimen telah cukup banyak dilakukan. Beberapa di antaranya adalah seperti penelitian (Turjaman & Budi, 2022) yang membagi sentimen berdasarkan teori bauran pemasaran dan menghasilkan bahwa aspek produk dan tempat didominasi sentimen negatif, aspek harga didominasi sentimen netral, dan aspek promosi didominasi sentimen positif. Penelitian (Naquitasia, 2022) membagi aspek berdasarkan aspek *mosque*, *halal food*, dan toilet. Hasil penelitian ini adalah aspek yang paling banyak dibahas adalah aspek *halal food* yang didominasi sentimen positif, kemudian aspek toilet didominasi sentimen negatif, dan aspek *mosque* jarang dibahas tetapi didominasi sentimen positif. Pada penelitian (Puspita & Widodo, 2021) didapatkan hasil bahwa tingkat akurasi metode *Decision Tree* sebesar 96,13%, lebih besar daripada metode KNN dengan akurasi 96,01% dan metode Naive Bayes dengan akurasi 89,14%.

Penelitian ini bertujuan untuk mengetahui bagaimana sentimen pengguna dan hasil implementasi metode *Decision Tree* terhadap faktor penggunaan PayPal berdasarkan teori bauran pemasaran 4P melalui ulasan pada Google Play Store. Penelitian ini diharapkan dapat menjadi referensi bagi pihak PayPal untuk meningkatkan kualitas layanan PayPal sehingga dapat bersaing dengan baik di Indonesia.

METODE

Proses umum analisis sentimen biasanya dimulai dari pengumpulan data baik secara *scraping* atau *crawling* sebuah situs web, kemudian dilakukan *preprocessing* seperti proses *tokenization*, *stop words removal*, *Part of Speech (POS) tagging*, dan *lemmatization*, selanjutnya dilakukan ekstraksi fitur dan klasifikasi sentimen, dan terakhir adalah proses analisis hasil yang didapatkan (Wankhade et al., 2022). Proses umum ABSA dapat dibagi menjadi 2 proses penting yaitu proses ekstraksi aspek yang menentukan aspek yang dibahas dalam suatu kalimat dan proses klasifikasi sentimen yang menentukan sentimen setiap aspek (Birjali et al., 2021).

Gambar 1 merupakan tahapan yang diterapkan pada penelitian ini. Tahapan penelitian dimulai dengan studi literatur mengenai analisis sentimen berbasis aspek terhadap dompet digital. Selanjutnya dilakukan ulasan Google Play Store dikumpulkan dengan menggunakan metode *scraping* data. *Scraping* situs web merupakan metode ekstraksi data tertentu secara otomatis dari situs web (A. Yani et al., 2019; Le, 2023). *Scraping* situs web ini dilakukan menggunakan bahasa pemrograman Python melalui Google Colaboratory. *Scraping* dilakukan dengan mengunduh *library* Google Play Scraper, lalu melakukan *import packages* yang diperlukan seperti *app*, *pandas*, *numpy*, *sort*, dan *reviews*. Data yang diambil ini merupakan data yang berbahasa Indonesia, berasal dari Indonesia, dan berjumlah 20.000 data.



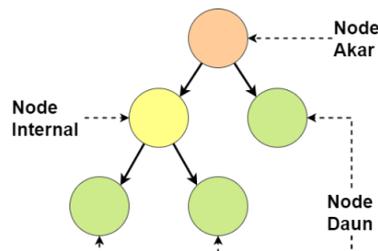
Gambar 1. Tahapan Penelitian

Tahapan *preprocessing* merupakan proses penyeleksian data setiap dokumen. Proses yang umum dilakukan pada *preprocessing* adalah *case folding*, *tokenizing*, *filtering*, *stop word removal*, dan *stemming* (Rosid et al., 2020; Widyassari et al., 2022). Pada penelitian ini, tahapan *preprocessing* yang diterapkan adalah *case folding*, *tokenizing*, *stop word removal*, dan *stemming*. *Case folding* merupakan proses yang mengubah seluruh huruf besar menjadi huruf kecil. *Tokenize* merupakan proses yang memisahkan kalimat menjadi token kata. *Stop word removal* merupakan proses yang menghapus kata-kata yang tidak diperlukan atau tidak bermakna (*stop word*). *Stemming* merupakan proses memotong awalan dan akhiran suatu kata sehingga kembali menjadi kata dasar (Hickman et al., 2022).

Klasifikasi aspek merupakan tahap yang dibentuk sesuai proses analisis sentimen berbasis aspek. Proses klasifikasi aspek terbagi menjadi proses *POS tagging*, pembuatan kamus kata, dan pelabelan aspek. *Part of Speech* (*POS tagging*) adalah proses yang menandai setiap kata sesuai jenis katanya, misalnya kata ‘suka’ dan ‘senang’ akan ditandai sebagai kata adjektif. Kata-kata yang telah ditandai selanjutnya dikelompokkan berdasarkan aspeknya dan dibuat kamus kata. Proses pelabelan aspek merupakan proses menandai setiap data ulasan sesuai aspeknya masing-masing berdasarkan kamus kata yang telah dibuat sebelumnya menggunakan Google Colaboratory.

Klasifikasi sentimen merupakan tahap yang disesuaikan dengan metode yang digunakan yaitu metode *Decision Tree*. Proses klasifikasi sentimen ini terbagi menjadi pelabelan sentimen, proses pemeriksaan label dan transformasi data, dan penerapan model *Decision Tree*. Pelabelan sentimen adalah proses yang menandai data sesuai sentimennya masing-masing. Proses ini dilakukan menggunakan Google Colaboratory dengan *library* *TextBlob*. Setelah data memiliki label aspek dan sentimen, data diperiksa secara manual dan dikelompokkan sesuai dengan aspek masing-masing. Proses ini akan menghasilkan 4 *dataset*, yaitu *dataset* produk, *dataset* harga, *dataset* tempat, dan *dataset* promosi. *Dataset-dataset* tersebut selanjutnya akan diuji menggunakan *Cross Validation* dan diterapkan model *Decision Tree*.

Decision Tree (pohon keputusan) merupakan pendekatan *machine learning* secara *supervised learning* yang digunakan untuk mengklasifikasikan polaritas sentimen dalam teks. Metode ini membangun struktur mirip pohon menggunakan data latih dengan tujuan membentuk model yang mampu memprediksi hasil klasifikasi atau regresi (Wankhade et al., 2022). *Decision Tree* memiliki beberapa kelebihan yaitu mampu mengelola data yang banyak, fleksibilitas yang baik, hasilnya mudah dipahami, dan akurasi yang lebih baik, tetapi jika ada kelas dan kriteria yang banyak akan lebih kompleks (Birjali et al., 2021; Fauziningrum & Suryaningsih, 2021; Puspita & Widodo, 2021; Sukerta Wijaya et al., 2021). Gambar 2 merupakan ilustrasi dari struktur *Decision Tree* yang terdiri dari *node* akar (*root node*), *node* internal (*internal node*), dan *node* daun (*leaf node*). *Node* akar merupakan *node* awal yang tidak memiliki data masukan. *Node* internal adalah *node* yang hanya memiliki satu data masukan dan memiliki data keluaran. Terakhir adalah *node* daun yaitu *node* yang hanya memiliki satu data masukan dan tidak memiliki data keluaran (Fitriani, 2020).



Gambar 2. Ilustrasi Decision Tree

Ada beberapa algoritma *Decision Tree* seperti algoritma C4.5, ID3, dan CART (Wankhade et al., 2022). Setiap algoritma memiliki kelebihan masing-masing, contohnya algoritma C4.5 yang merupakan pengembangan lebih lanjut dari algoritma ID3, lebih baik dalam mengatasi data yang hilang, data yang bersifat kontinyu, dan dapat melakukan *pruning*. Tahapan membentuk *Decision Tree* dengan algoritma C4.5 secara umum dimulai dengan pemilihan atribut untuk menjadi node akar. Selanjutnya dibentuk cabang pada setiap nilai. Setelah cabang dibuat, kasus dibagi pada cabang. Proses terakhir adalah pengulangan proses pada cabang sampai dihasilkan kelas yang sama pada semua kasus cabang (Ma'ruf et al., 2021).

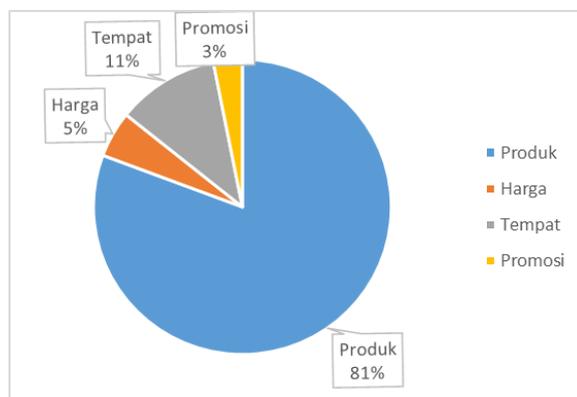
Penerapan model *Decision Tree* pada aplikasi RapidMiner menggunakan kriteria *gain ratio*, *information gain*, dan *gini index*. Pada proses *Cross Validation*, nilai k yang digunakan k=2, k=3, k=4, dan k=5. Proses ini menghasilkan perhitungan kinerja evaluasi model *Decision Tree* seperti akurasi, presisi, dan *recall*. Setelah proses-proses tersebut dilakukan, maka akan dilakukan analisis hasil dari klasifikasi sentimen. Proses ini mengidentifikasi nilai *fold* dan kriteria *Decision Tree* yang memiliki akurasi terbaik secara keseluruhan. Proses ini juga akan menjelaskan bagaimana sentimen masyarakat Indonesia terhadap aspek penggunaan PayPal.

Data hasil dari penerapan *Decision Tree* dan pengujian *K-Fold Cross Validation* selanjutnya akan dianalisis sehingga dapat diketahui informasi apa saja yang dihasilkan. Dari hasil analisis ini kemudian dibentuk suatu kesimpulan dan saran untuk penelitian selanjutnya.

HASIL DAN PEMBAHASAN

Analisis Pembagian Data

Data yang dikumpulkan dari Google Play Store sebesar 20.000 data. Data hasil *pre-processing* berjumlah 17.012 data. Proses klasifikasi aspek menghasilkan 12.324 data terlabel, yang mencakup 9.944 data produk, 623 data harga, 1.371 data tempat, dan 386 data promosi. Pada tabel 1 dan gambar 3, dapat diperhatikan bahwa 81% pengguna cenderung memberikan ulasan mengenai aspek produk PayPal, sementara 11% memberikan ulasan mengenai aspek tempat, 5% memberikan ulasan mengenai aspek harga, dan hanya 3% pengguna yang memberikan ulasan mengenai aspek promosi. Analisis ini mengindikasikan bahwa pengguna PayPal lebih sering mengulas mengenai aspek produk dibandingkan dengan aspek lainnya, dan aspek promosi PayPal lebih jarang menjadi fokus ulasan dibandingkan dengan aspek lainnya. Berikut ini adalah tabel dan grafik pembagian data berdasarkan aspek.



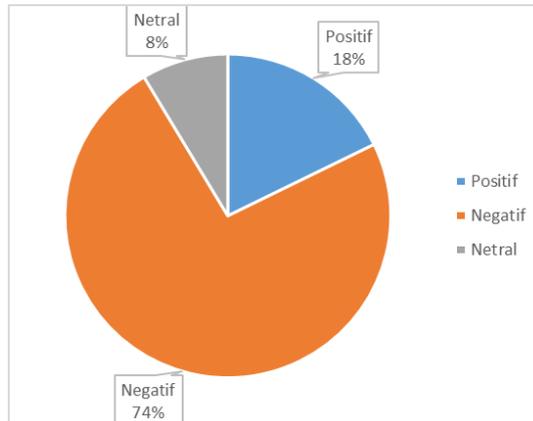
Gambar 3. Grafik Pembagian Data berdasarkan Aspek

Tabel 1. Pembagian Data berdasarkan Aspek

Aspek	Jumlah Data
Produk	9.944
Harga	623
Tempat	1.371
Promosi	386
Total Data	12.324

Proses klasifikasi sentimen seperti pada tabel 2 dan gambar 4 menghasilkan 12.324 data, terdiri dari 2.184 data positif, 9.081 data negatif, dan 1.059 data netral. Dari pembagian data berdasarkan sentimen ini, dapat disimpulkan bahwa pengguna PayPal memberikan ulasan dengan 74% sentimen negatif, 18% sentimen positif, dan 8% sentimen netral. Analisis ini menunjukkan bahwa pengguna PayPal lebih sering menyampaikan

ulasan dengan sentimen negatif dibandingkan dengan ulasan positif atau netral. Berikut ini adalah tabel dan grafik pembagian data berdasarkan sentimen.



Gambar 4. Grafik Pembagian Data Berdasarkan Sentimen

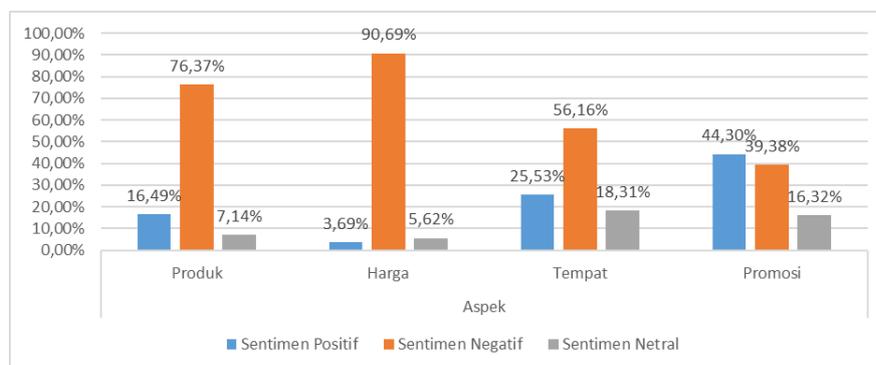
Tabel 2. Pembagian Data Berdasarkan Sentimen

Sentimen	Jumlah Data
Positif	2.184
Negatif	9.081
Netral	1.059
Total Data	12.324

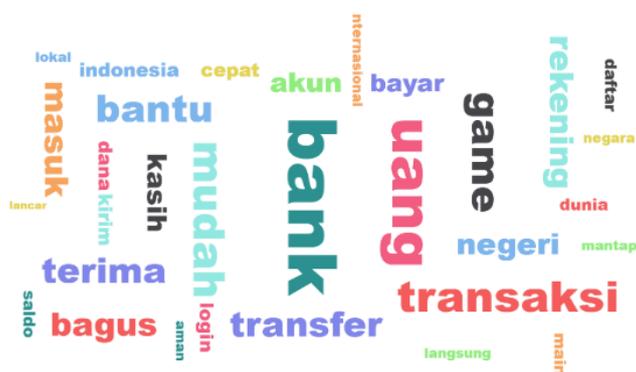
Secara keseluruhan seperti pada tabel 3 dan gambar 5, data positif pada aspek produk berjumlah 1.640, data negatif berjumlah 7.594, dan data netral berjumlah 710. Untuk aspek harga, data positif berjumlah 23, data negatif berjumlah 565, dan data netral berjumlah 35. Pada aspek tempat, data positif berjumlah 350, data negatif berjumlah 770, dan data netral berjumlah 251. Sementara itu, pada aspek promosi, data positif berjumlah 171, data negatif berjumlah 152, dan data netral berjumlah 63. Dari pembagian data berdasarkan aspek dan sentimen ini, terlihat bahwa pada aspek produk, 76% pengguna PayPal memberikan sentimen negatif, 17% sentimen positif, dan 7% sentimen netral. Pada aspek harga, pengguna memberikan 91% sentimen negatif, 6% sentimen netral, dan 4% sentimen positif. Untuk aspek tempat, pengguna memberikan 56% sentimen negatif, 26% sentimen positif, dan 18% sentimen netral. Pada aspek promosi, pengguna memberikan 44% sentimen positif, 39% sentimen negatif, dan 16% sentimen netral.

Tabel 3. Pembagian Data Berdasarkan Aspek dan Sentimen

Sentimen	Produk	Harga	Tempat	Promosi
Positif	1.640	23	350	171
Negatif	7.594	565	770	152
Netral	710	35	251	63
Total	9.944	623	1.371	386



Gambar 5. Grafik Pembagian Data Berdasarkan Aspek dan Sentimen



Gambar 10. Wordcloud Aspek Tempat Sentimen Positif



Gambar 11. Wordcloud Aspek Tempat Sentimen Negatif

Berdasarkan Gambar 12, dapat disimpulkan bahwa kata yang paling sering disebut oleh pengguna PayPal pada aspek promosi dengan sentimen positif adalah ‘hadiah’ yang muncul sebanyak 62 kali. Kata ‘kasih’ (42 kali), ‘bonus’ (39 kali), dan ‘gratis’ (33 kali) menunjukkan bahwa unsur insentif atau keuntungan tambahan yang diterima oleh pengguna. Ungkapan rasa terima kasih juga terlihat dalam kata-kata ‘terima’ (27 kali) dan ‘terimakasih’ (27 kali). Beberapa kata seperti ‘dapat’ (25 kali), ‘saldo’ (22 kali), ‘uang’ (22 kali), dan ‘dolar’ (18 kali) menunjukkan keterkaitan dengan bentuk hadiah uang yang diberikan PayPal. Ungkapan positif seperti ‘mantap’ (16 kali), ‘moga’ (16 kali), dan ‘bagus’ (14 kali) menandakan evaluasi positif terhadap aspek promosi. Analisis kata-kata ini menunjukkan bahwa pengguna PayPal memiliki pandangan positif terhadap aspek promosi PayPal karena puas akan hadiah yang diberikan PayPal pada pengguna.



Gambar 12. Wordcloud Aspek Promosi Sentimen Positif

Berdasarkan Gambar 13, dapat disimpulkan bahwa kata yang paling sering disebut oleh pengguna PayPal pada aspek promosi dengan sentimen negatif adalah ‘hadiah’ yang disebutkan sebanyak 74 kali. Kata ‘bonus’ disebutkan 34 kali dan ‘saldo’ sebanyak 29 kali. Ungkapan seperti ‘dapat’ (28 kali), ‘masuk’ (25 kali), dan ‘kasih’ (22 kali) memberikan gambaran pengalaman pengguna dalam menerima hadiah atau bonus.

Beberapa kata seperti 'claim' (18 kali) dan 'tipu' (18 kali) menunjukkan ketidakpuasan atau kecurigaan terhadap klaim atau tawaran. Analisis kata-kata ini menunjukkan bahwa pengguna PayPal memiliki pandangan negatif terhadap aspek promosi PayPal karena beberapa pengguna tidak mendapat hadiah meskipun telah setia menggunakan PayPal dan beberapa pengguna lain merasa tertipu atas hadiah yang diberikan.



Gambar 13. Wordcloud Aspek Promosi Sentimen Negatif

Analisis Hasil Pengujian K-Fold Cross Validation

Penerapan model *Decision Tree* dilakukan dengan kriteria *gain ratio*, *information gain*, dan *gini index* dengan nilai pengujian *Cross Validation* $k=2$, $k=3$, $k=4$, dan $k=5$. Pada tabel 12, dapat diidentifikasi bahwa dalam pengujian *fold* ke-2, nilai akurasi tertinggi terdapat pada aspek harga dengan kriteria *gain ratio* mencapai 89,73%, presisi tertinggi ada pada aspek produk dengan kriteria *gini index* sebesar 65,89%, serta *recall* tertinggi ditemukan pada aspek promosi dengan kriteria *gini index* sebesar 55,05%. Pada pengujian *fold* ke-3, akurasi tertinggi kembali dicapai oleh aspek harga kriteria *gain ratio* sebesar 90,37%, presisi tertinggi terdapat pada aspek produk dengan kriteria *information gain* sebesar 67,38%, dan *recall* tertinggi pada aspek promosi dengan kriteria *gini index* sebesar 55,62%. Pada pengujian *fold* ke-4, nilai akurasi tertinggi pada aspek harga dengan kriteria *gain ratio* sebesar 89,73%, presisi tertinggi pada aspek produk dengan kriteria *gain ratio* sebesar 71,49%, dan *recall* tertinggi pada aspek promosi dengan kriteria *gini index* sebesar 53,95%. Terakhir, pada pengujian *fold* ke-5, akurasi tertinggi dicapai oleh aspek harga dengan kriteria *gain ratio* sebesar 89,73%, presisi tertinggi pada aspek produk dengan kriteria *gain ratio* sebesar 70,85%, dan *recall* tertinggi pada aspek promosi dengan kriteria *gini index* sebesar 54,73%.

Berdasarkan pengamatan tersebut, terlihat bahwa akurasi tertinggi dari nilai *fold* ke-2 hingga nilai *fold* ke-5 diperoleh oleh aspek harga dengan kriteria *gain ratio*, dan akurasi tertinggi terdapat pada nilai *fold* ke-5, mencapai 90,69%. Presisi tertinggi ditemukan pada nilai *fold* ke-4 oleh aspek produk dengan kriteria *gain ratio*, mencapai 71,49%. Sementara itu, *recall* tertinggi diperoleh pada nilai *fold* ke-3 oleh aspek promosi dengan kriteria *gini index*, mencapai 55,62%. Oleh sebab itu, dapat disimpulkan bahwa kriteria *gain ratio* menghasilkan model yang mampu mengidentifikasi klasifikasi sentimen secara akurat, dengan akurasi tertinggi mencapai 90,69%. Kriteria *gain ratio* juga efektif dalam menghasilkan model yang dapat mengurangi kesalahan prediksi, dengan nilai tertinggi presisi sebesar 71,49%. Kriteria *gini index*, sementara itu, menghasilkan model yang mampu mengenali sebagian besar label atau sentimen yang benar, dengan nilai tertinggi *recall* sebesar 55,62%.

Dari hasil pengamatan tabel 4, dapat dibuat suatu perhitungan untuk menentukan nilai k yang memiliki hasil akurasi terbaik untuk kriteria *gain ratio* dalam pengujian *Decision Tree* seperti yang terlihat pada perhitungan tabel 5. Pada tabel 5, ditentukan nilai skor berdasarkan jumlah akurasinya sesuai aspek masing-masing. Misalnya pada aspek produk, pada nilai akurasi tertinggi dicapai pada 4-Fold Cross Validation, sehingga nilai $k=4$ diberi skor 4 poin. Langkah selanjutnya adalah menghitung total skor akurasi berdasarkan nilai k masing-masing. Dari perhitungan ini diketahui bahwa nilai $k=5$ menghasilkan total skor sebesar 12 poin yang merupakan skor terbesar dibandingkan nilai k lainnya, hal ini berarti hasil akurasi nilai $k=5$ lebih baik secara keseluruhan. Oleh sebab itu, dalam penelitian ini, 5-Fold Cross Validation lebih cocok digunakan.

Nilai metrik yang didapatkan berdasarkan kriteria *gain ratio* dan $k=5$ adalah nilai akurasi 78,34%, presisi 70,85%, dan *recall* 38,00% untuk aspek produk. Pada aspek harga, nilai akurasi 90,69%, presisi 46,94%, dan *recall* 34,72%. Pada aspek tempat, nilai akurasi 57,40%, presisi 60,46%, dan *recall* 36,00%. Pada aspek promosi, nilai akurasi 68,65%, presisi 63,33%, dan *recall* 54,62%.

Tabel 4. Hasil Pengujian *Decision Tree* Seluruh Aspek

K	Metrik Evaluasi	Produk			Harga			Tempat			Promosi		
		<i>Gain ratio</i>	<i>Information gain</i>	<i>Gini index</i>	<i>Gain ratio</i>	<i>Information gain</i>	<i>Gini index</i>	<i>Gain ratio</i>	<i>Information gain</i>	<i>Gini index</i>	<i>Gain ratio</i>	<i>Information gain</i>	<i>Gini index</i>
2	Akurasi	77,00%	84,14%	84,16%	89,73%	89,08%	85,87%	57,91%	60,83%	64,48%	56,48%	63,99%	67,10%
	Presisi	63,36%	63,14%	65,89%	38,67%	47,37%	45,37%	52,65%	46,47%	57,35%	51,36%	52,82%	55,48%
	Recall	35,14%	52,05%	52,01%	35,76%	40,49%	39,74%	36,70%	43,31%	48,61%	44,59%	52,42%	55,05%
3	Akurasi	78,91%	83,89%	84,36%	90,37%	89,40%	87,32%	56,97%	62,22%	63,02%	69,17%	66,84%	66,85%
	Presisi	67,30%	67,88%	66,17%	30,27%	45,59%	32,44%	59,43%	47,50%	53,28%	52,97%	53,57%	56,03%
	Recall	39,27%	51,57%	52,41%	33,22%	39,32%	33,49%	34,97%	48,66%	47,14%	55,06%	54,12%	55,62%
4	Akurasi	79,57%	83,85%	84,34%	89,89%	88,45%	89,09%	56,60%	62,73%	64,63%	66,57%	65,03%	66,83%
	Presisi	71,49%	67,22%	68,78%	35,01%	35,45%	45,45%	51,59%	53,17%	56,30%	45,57%	51,37%	52,27%
	Recall	40,63%	51,52%	52,92%	34,43%	34,21%	39,20%	35,17%	46,69%	48,83%	52,61%	52,92%	53,95%
5	Akurasi	78,34%	83,30%	84,11%	90,69%	89,09%	87,96%	57,40%	62,73%	63,02%	68,65%	67,36%	67,89%
	Presisi	70,85%	59,93%	66,37%	46,94%	51,77%	40,50%	60,46%	48,82%	50,12%	63,33%	53,07%	52,97%
	Recall	38,00%	50,07%	52,36%	34,72%	39,20%	37,39%	36,00%	48,12%	48,62%	54,62%	54,29%	54,73%

Tabel 5. Perhitungan Nilai K Terbaik Kriteria *Gain Ratio*

Nilai K	Produk		Harga		Tempat		Promosi		Total Skor
	Akurasi	Skor	Akurasi	Skor	Akurasi	Skor	Akurasi	Skor	
2	77,00%	1	89,73%	1	57,91%	4	56,48%	1	7
3	78,91%	3	90,37%	3	56,97%	1	69,17%	4	11
4	79,57%	4	89,89%	2	56,60%	2	66,57%	2	10
5	78,34%	2	90,69%	4	57,40%	3	68,65%	3	12

Keterangan:

Akurasi tertinggi pertama = skor 4 poin

Akurasi tertinggi kedua = skor 3 poin

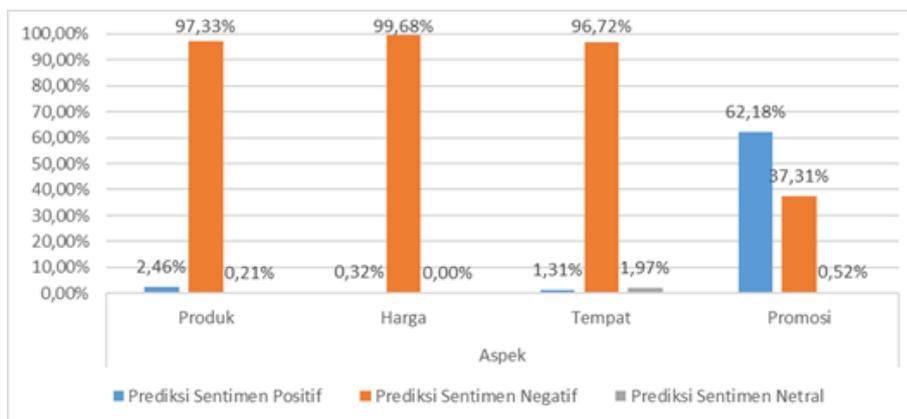
Akurasi tertinggi ketiga = skor 2 poin

Akurasi tertinggi keempat = skor 1 poin

Tabel 6 dan gambar 14 menampilkan hasil prediksi sentimen terhadap berbagai aspek penggunaan PayPal dengan menggunakan kriteria *gain ratio* pada *fold* ke-5. Secara keseluruhan, prediksi terhadap data produk menunjukkan bahwa pengguna aplikasi PayPal memberikan ulasan dengan 97,33% sentimen negatif, 2,46% sentimen positif, dan 0,21% sentimen netral. Pada aspek harga, prediksi menunjukkan bahwa pengguna memberikan ulasan dengan 99,68% sentimen negatif, 0,32% sentimen positif, dan 0% sentimen netral. Pada aspek tempat, pengguna memberikan ulasan dengan 96,72% sentimen negatif, 1,97% sentimen netral, dan 1,31% sentimen positif. Terakhir, pada aspek promosi, prediksi menunjukkan bahwa pengguna memberikan ulasan dengan 62,18% sentimen positif, 37,31% sentimen negatif, dan 0,82% sentimen netral. Dari analisis ini, dapat disimpulkan bahwa pengguna PayPal masih cenderung memberikan sentimen negatif terutama pada aspek produk, harga, dan tempat PayPal, memberikan peluang untuk peningkatan kualitas di ketiga aspek tersebut. Sebaliknya, aspek promosi mendapatkan sentimen positif, menandakan bahwa kualitas layanan promosi PayPal dianggap cukup baik.

Tabel 6. Hasil Prediksi Sentimen 5-Fold Cross Validation Kriteria Gain Ratio

Sentimen	Produk	Harga	Tempat	Promosi
Positif	245	2	18	240
Negatif	9.678	621	1.326	144
Netral	21	0	27	2
Total	9.944	623	1.371	386



Gambar 14. Grafik Hasil Prediksi Sentimen 5-Fold Cross Validation Kriteria Gain Ratio

KESIMPULAN DAN SARAN

Kesimpulan yang didapat dari hasil analisis penerapan model *Decision Tree* adalah pengguna PayPal cenderung memberikan sentimen negatif pada aspek produk (97,33% sentimen negatif), harga (99,68% sentimen negatif), dan tempat (96,72% sentimen negatif). Hal ini membuka peluang bagi PayPal untuk melakukan kegiatan peningkatan kualitas di ketiga aspek tersebut. Sebaliknya, pada aspek promosi didominasi oleh sentimen positif (62,18%), menandakan bahwa kualitas layanan promosi PayPal telah dianggap cukup baik sehingga perlu dipertahankan.

Berdasarkan analisis *wordcloud*, evaluasi negatif terhadap aspek produk terfokus pada fitur pendaftaran akun, login akun, dan transaksi keuangan karena dianggap sulit dan seringkali mengalami kesalahan. Pada aspek harga, pandangan negatif muncul karena pengguna tidak puas terhadap pemotongan biaya transaksi dan biaya PayPal lainnya. Aspek tempat mendapat kritik negatif karena pengguna tidak dapat terhubung pada tempat PayPal seharusnya dapat digunakan terutama pada aktivitas permainan dalam jaringan. Sebaliknya, aspek promosi dinilai positif oleh pengguna PayPal karena pengguna puas akan hadiah yang diberikan PayPal.

Walaupun penelitian mengenai analisis sentimen faktor penggunaan PayPal telah selesai dilakukan, masih ada peluang untuk penelitian lanjutan. Peluang tersebut misalnya seperti penggunaan objek dan metode pengujian yang lebih terbaru, teknik pelabelan yang lebih baik, data yang lebih seimbang, dan aspek lain yang terkait objek penelitian. Dengan demikian, penelitian mengenai analisis sentimen ini dapat menjadi referensi untuk penelitian-penelitian selanjutnya.

DAFTAR RUJUKAN

A. Yani, D. D., Pratiwi, H. S., & Muhandi, H. (2019). Implementasi Web Scraping untuk Pengambilan Data pada Situs Marketplace. *Jurnal Sistem Dan Teknologi Informasi (JUSTIN)*, 7(4), 257. <https://doi.org/10.26418/justin.v7i4.30930>

Abdullah, T. (2019). *Manajemen Pemasaran* (Cetakan 1). Rajawali Pers.

Arbaini, P. (2020). Pengaruh Consumer Online Rating Dan Review Terhadap Keputusan Pembelian Pada Pengguna Marketplace Tokopedia. *Jurnal Bisnis Dan Manajemen*, 7(1), 25–33. <https://doi.org/10.26905/jbm.v7i1.3897>

Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, 226, 107134. <https://doi.org/10.1016/j.knosys.2021.107134>

Daryanto, L. H., Hasiholan, L. B., & Seputro, A. (2019). The Influence Of Marketing Mix On The Decision To Purchase Martabak ‘Setiabudi’ Pak Man Semarang. *Journal of Management*, 5(5), 1–7.

Fauziningrum, E., & Suryaningsih, E. I. (2021). Evaluasi Dan Prediksi Penguasaan Bahasa Inggris Maritim Menggunakan Metode Decision Tree Dan Confusion Matrix (Studi Kasus Di Universitas Maritim

- Amni). *Angewandte Chemie International Edition*, 6(11), 951–952., 5–24.
- Fitriani, E. (2020). Perbandingan Algoritma C4.5 dan Naive Bayes untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan. *SISTEMASI*, 9(1), 103. <https://doi.org/10.32520/stmsi.v9i1.596>
- Fuadah, Y. N., Ubaidullah, I. D., Ibrahim, N., Taliningsing, F. F., Sy, N. K., & Pramuditho, M. A. (2022). Optimasi Convolutional Neural Network dan K-Fold Cross Validation pada Sistem Klasifikasi Glaukoma. *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, 10(3), 728. <https://doi.org/10.26760/elkomika.v10i3.728>
- Hapsari, R. A., Hartono, B., & Listiyani, Y. (2021). Dispute Settlement Analysis and Refused Claims in Payment Gateway Transactions on Paypal Indonesia. *Lambung Mangkurat Law Journal*, 6(2), 199–212. <https://doi.org/10.32801/lamlaj.v6i2.262>
- Hickman, L., Thapa, S., Tay, L., Cao, M., & Srinivasan, P. (2022). Text Preprocessing for Text Mining in Organizational Research: Review and Recommendations. *Organizational Research Methods*, 25(1), 114–146. <https://doi.org/10.1177/1094428120971683>
- Hoang, M., Alija Bihorac, O., & Rouces, J. (2019). Aspect-Based Sentiment Analysis Using BERT. *Proceedings of the 22nd Nordic Conference on Computational Linguistics*, 187–196.
- Le, S. (2023). *Web Scraping of University Rankings and Data Analysis Using Python*.
- Lestari, W., Musyahidah, S., & Istiqamah, R. (2019). Strategi Marketing Mix Dalam Meningkatkan Usaha Percetakan Pada CV. Tinta Kaili dalam Perspektif Ekonomi Islam. *Jurnal Ilmu Ekonomi Dan Bisnis Islam*, 1(1), 63–84. <https://doi.org/10.24239/jiebi.v1i1.5.63-84>
- Ma'ruf, F. A., Pratama, A., Sholihin, I., & Rizki Rinaldi, A. (2021). Penerapan Model Prediksi Menggunakan Algoritma C.45 Untuk Prediksi Kelulusan Siswa SMK Wahidin. *Jurnal Data Science & Informatika*, 1(1), 16–20.
- Majidah, S., & Istianah, I. (2023). Analisis Strategi Bauran Pemasaran 4P di Era 5.0 Berbasis Maqāsid Syarīah. *Jurnal Hukum Ekonomi Syariah*, 6(1), 59. <https://doi.org/10.30595/jhes.v6i1.14690>
- Naquitasia, R. (2022). *Analisis Sentimen Berbasis Aspek Pada Ulasan Tempat Wisata Halal Dengan Deep Learning* [Universitas Islam Indonesia]. <https://dspace.uui.ac.id/handle/123456789/39202%0Ahttps://dspace.uui.ac.id/bitstream/handle/123456789/39202/18523214.pdf?sequence=1>
- Pamungkas, T. J., & Romadhony, A. (2021). Analisis Sentimen Berbasis Aspek Terhadap Ulasan Restoran Berbahasa Indonesia menggunakan Support Vector Machines. *E-Proceeding of Engineering*, 8(4), 4102–4114.
- Permana, A. P., Ainayah, K., & Holle, K. F. H. (2021). Analisis Perbandingan Algoritma Decision Tree, KNN, dan Naive Bayes untuk Prediksi Kesuksesan Start-up. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 6(3), 178–188. <https://doi.org/10.14421/jiska.2021.6.3.178-188>
- Peryanto, A., Yudhana, A., & Umar, R. (2020). Klasifikasi Citra Menggunakan Convolutional Neural Network dan K Fold Cross Validation. *Journal of Applied Informatics and Computing*, 4(1), 45–51. <https://doi.org/10.30871/jaic.v4i1.2017>
- Prasasti, G. D. (2022, December 6). Survei Jakpat: DANA Jadi Dompot Digital Terpopuler di Paruh Pertama 2022. *Liputan 6*. [Online]. <https://www.liputan6.com/tekno/read/5145618/survei-jakpat-dana-jadi-dompot-digital-terpopuler-di-paruh-pertama-2022>
- Puspita, R., & Widodo, A. (2021). Perbandingan Metode KNN, Decision Tree, dan Naive Bayes Terhadap Analisis Sentimen Pengguna Layanan BPJS. *Jurnal Informatika Universitas Pamulang*, 5(4), 646. <https://doi.org/10.32493/informatika.v5i4.7622>
- Respati, A. R. (2022, November 28). Survei InsightAsia: 71 Persen Masyarakat Gunakan Dompot Digital, GoPay di Posisi Pertama. *Kontan.Co.Id*. [Online]. <https://money.kompas.com/read/2022/11/28/164000226/survei-insightasia-71-persen-masyarakat-gunakan-dompot-digital-gopay-di-posisi>
- Rosid, M. A., Fitriani, A. S., Astutik, I. R. I., Mulloh, N. I., & Gozali, H. A. (2020). Improving Text Preprocessing For Student Complaint Document Classification Using Sastrawi. *IOP Conference Series: Materials Science and Engineering*, 874(1), 012017. <https://doi.org/10.1088/1757-899X/874/1/012017>
- Saputra, S. A., Rosiyadi, D., Gata, W., & Husain, S. M. (2019). Google Play E-Wallet Sentiment Analysis Using Naive Bayes Algorithm Based on Particle Swarm Optimization. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 3(3), 377–382.
- Sari, F. V., & Wibowo, A. (2019). Analisis Sentimen Pelanggan Toko Online Jd.Id Menggunakan Metode Naive Bayes Classifier Berbasis Konversi Ikon Emosi. *Jurnal SIMETRIS*, 10(2), 681–686.
- Sepfiani, P., & Nawawi, Z. M. (2023). *Analisis Marketing Mix untuk Meningkatkan Volume Penjualan pada*

- Rafijaya Fotokopi di Kota Medan*. 2(1), 56–62.
- Sukerta Wijaya, I. W., Harjumawan Wiratmaja KS., I. G., Pramana Setya Bintara, I. D. M. A., & Ryan Aditya Permana, I. K. G. (2021). Program Menghitung Banyak Bata pada Ruangan Menggunakan Bahasa Python. *TIERS Information Technology Journal*, 2(1). <https://doi.org/10.38043/tiers.v2i1.2840>
- Turjaman, R. M., & Budi, I. (2022). Analisis Sentimen Berbasis Aspek Marketing Mix Terhadap Ulasan Aplikasi Dompot Digital (Studi Kasus: Aplikasi Linkaja Pada Twitter). *Jurnal Darma Agung*, 30(2), 266. <https://doi.org/10.46930/ojsuda.v30i2.1672>
- Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7), 5731–5780. <https://doi.org/10.1007/s10462-022-10144-1>
- Widi, S. (2022). Jumlah Pengguna Paypal Capai 429 Juta per Kuartal I/2022. *Dataindonesia.Id. [Online]*. <https://dataindonesia.id/digital/detail/jumlah-pengguna-paypal-capai-429-juta-per-kuartal-i2022>
- Widyassari, A. P., Rustad, S., Shidik, G. F., Noersasongko, E., Syukur, A., Affandy, A., & Setiadi, D. R. I. M. (2022). Review of automatic text summarization techniques & methods. *Journal of King Saud University - Computer and Information Sciences*, 34(4), 1029–1046. <https://doi.org/10.1016/j.jksuci.2020.05.006>
- Wisudawati, D. T., Utami, T. iani W., & Arum, P. R. (2021). Analisis Sentimen Terhadap Dampak Covid-19 Pada Performa Tokopedia Menggunakan Support Vector Machine. *Seminar Nasional Variansi ...*, 87–96. <https://ojs.unm.ac.id/variانسistatistika/article/view/19508>
- Zahrah, A., Mandey, S. L., & Mangantar, M. (2021). Analisis Marketing Mix Terhadap Volume Penjualan Pada UMKM RM. Solideo Kawasan Bahu Mall Manado. *Jurnal EMBA: Jurnal Riset Ekonomi, Manajemen, Bisnis Dan Akuntansi*, 9(4), 216–226. <https://ejournal.unsrat.ac.id/index.php/emba/article/view/36216>