
Implementasi *Naïve Bayes Classifier* untuk Sentimen Produk Kecantikan Berdasarkan Ulasan *Female Daily*

Pramesti Melinea Berlianti¹, Erwin Yudi Hidayat²

111202113528@mhs.dinus.ac.id¹, erwin@dsn.dinus.ac.id²

^{1,2} Universitas Dian Nuswantoro

Informasi Artikel

Diterima : 12 Nov 2024

Direvisi : 25 Nov 2024

Disetujui : 30 Dec 2024

Kata Kunci

Analisis Sentimen, ,
Female Daily, Klasifikasi,
Naïve Bayes Classifier

Abstrak

Industri kecantikan di Indonesia mengalami pertumbuhan yang pesat, terlihat dari bertambahnya perusahaan kecantikan setiap tahun. Salah satu produk kecantikan yang banyak diminati adalah *moisturizer*, yang menghasilkan banyak ulasan di platform kecantikan, seperti *Female Daily*. Penelitian ini bertujuan menganalisis sentimen ulasan pengguna terhadap produk kecantikan untuk memahami preferensi konsumen, menggunakan metode *Naïve Bayes Classifier*. Sebanyak 1.050 ulasan diolah melalui proses *preprocessing* hingga menghasilkan 1.042 data bersih, terdiri dari sentimen positif (84.36%), negatif (12.96%), dan netral (2.69%). Untuk menangani ketidakseimbangan data diterapkan teknik *SMOTE*, dan pencarian *hyperparameter* optimal dilakukan dengan *GridSearchCV*, yang meningkatkan akurasi model dari 79% menjadi 82,77%. Hasil ini menunjukkan bahwa penerapan *GridSearchCV* berperan penting dalam meningkatkan akurasi dan stabilitas klasifikasi sentimen pada produk kecantikan.

Keywords

Classification, Female Daily, Naive Bayes Classifier, Sentiment Analysis

Abstract

The beauty industry in Indonesia is experiencing rapid growth, as seen from the increase in beauty companies every year. One of the most popular beauty products is moisturizer, which generates many reviews on beauty platforms, such as Female Daily. This research aims to analyze the sentiment of user reviews on beauty products to understand consumer preferences, using the Naïve Bayes Classifier method. A total of 1,050 reviews were processed through preprocessing to produce 1,042 clean data, consisting of positive (84.36%), negative (12.96%), and neutral (2.69%) sentiments. To handle data imbalance, the SMOTE technique was applied, and the optimal hyperparameter search was performed with GridSearchCV, which improved the model accuracy from 79% to 82.77%. These results show that the application of GridSearchCV plays an important role in improving the accuracy and stability of sentiment classification on beauty product.

A. Pendahuluan

Industri kecantikan di Indonesia mengalami pertumbuhan yang pesat setiap tahunnya. Hal ini tercermin dari peningkatan jumlah perusahaan kecantikan sebesar 21,9% dibandingkan tahun 2022, yang tercatat sebanyak 913 perusahaan, dan meningkat menjadi 1.010 perusahaan pada tahun 2023, atau meningkat 10 kali lipat dalam 5 tahun terakhir [1].

Produk kecantikan dan perawatan kulit dipandang sebagai hal yang sangat penting untuk menjaga penampilan oleh banyak wanita [2]. Salah satu produk kecantikan yang paling penting untuk perawatan kulit adalah *moisturizer*, yang berguna menjaga kelembaban kulit agar tetap sehat dan terawat. Tingginya minat konsumen terhadap produk ini memicu banyaknya ulasan dan diskusi di *platform* kecantikan seperti *Female Daily* yang menjadi wadah utama untuk berbagi pengalaman pengguna tentang produk kecantikan yang dimaksud.

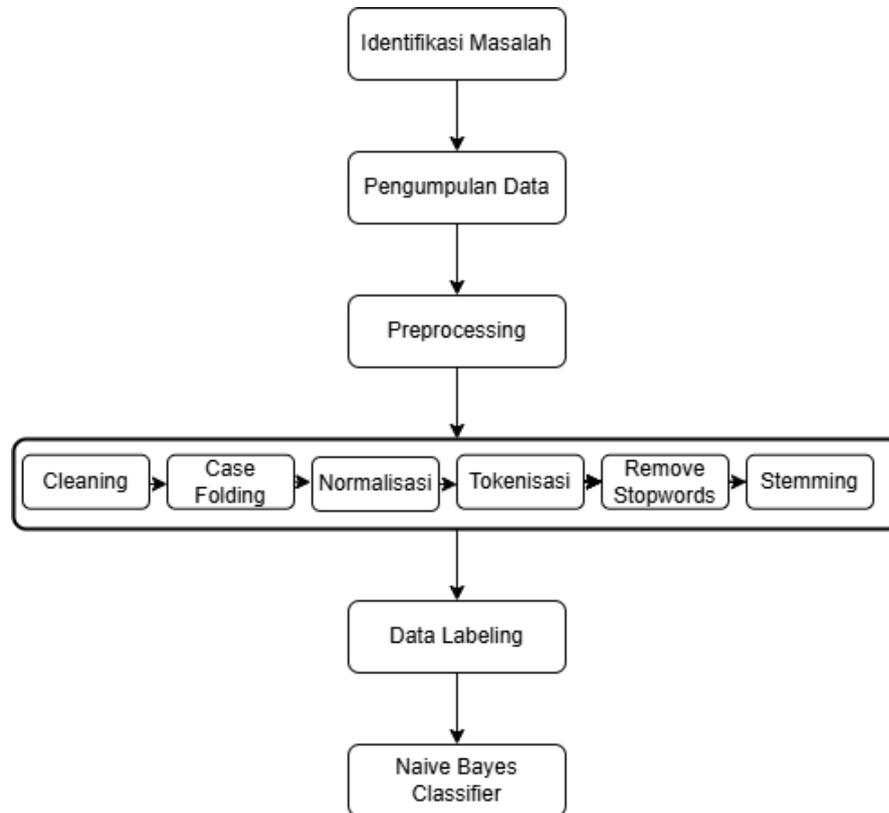
Analisis sentimen telah menjadi alat penting untuk memahami opini dan persepsi konsumen terhadap produk di berbagai industri, termasuk industri kecantikan. Analisis sentimen digunakan untuk mengidentifikasi bagaimana suatu sentimen diekspresikan dalam teks dan bagaimana sentimen dapat dikategorikan sebagai respon positif atau negatif [3]. Seiring dengan berkembangnya minat masyarakat terhadap produk kecantikan, analisis sentimen menjadi alat yang penting untuk memahami opini publik terhadap produk-produk tersebut.

Platform Female Daily sebagai komunitas pertama di Indonesia yang berfokus pada bidang kecantikan dipilih karena kredibilitasnya yang menyediakan ulasan terpercaya mengenai perawatan kulit, tubuh, dan rambut, serta kebutuhan wanita lainnya [4]. Informasi tentang produk dan layanan menjadi lebih valid dan terpercaya berkat ulasan dari pengguna sebelumnya. Rekomendasi dan komentar pengguna sebelumnya tentang produk dan layanan dipercaya oleh banyak pengguna media sosial [5]. Ulasan pengguna sering menjadi acuan konsumen saat memilih produk kecantikan. Dalam konteks ini, penelitian untuk menganalisis sentimen pengguna terhadap produk kecantikan sangat relevan, terutama mengingat tingginya keterlibatan pengguna dalam memberikan *review* pada produk kecantikan di *platform* tersebut.

Metode *Naïve Bayes* merupakan metode pembelajaran mesin yang digunakan untuk masalah klasifikasi, terutama dalam klasifikasi teks yang melibatkan kumpulan data pelatihan yang cukup besar. Contoh penggunaannya termasuk analisis sentimen, penyaringan spam, dan klasifikasi teks [6]. Metode ini dipilih karena memiliki waktu proses yang lebih cepat daripada algoritma pembelajaran mesin lainnya dan dikenal memiliki keakuratan prediksi yang tinggi [7]. Berbagai penelitian telah dilakukan dengan menggunakan metode *Naïve Bayes Classifier*, seperti penelitian [8] menghasilkan akurasi sebesar 79,85%. Pada penelitian [9] menghasilkan akurasi sebesar 80%. Berikutnya pada [10] menghasilkan akurasi sebesar 86,6%. Dengan demikian, penggunaan metode *Naïve Bayes Classifier* dapat mendukung pengambilan keputusan, memberikan pandangan yang lebih akurat mengenai preferensi konsumen, dan berkontribusi pada penerapan analisis sentimen.

B. Metode Penelitian

Penelitian ini menggunakan metode yang bertahap yang bertujuan untuk memberikan langkah sistematis dalam mencapai tujuan penelitian secara komprehensif. Prosedur penelitian ini dijelaskan sebagai berikut:



Gambar 1. Metodologi Penelitian

Prosedur yang dilakukan dapat diuraikan sebagai berikut:

1. **Identifikasi Masalah**
Pada tahap identifikasi masalah, penelusuran dilakukan untuk memahami konteks penelitian. Penelitian ini bertujuan untuk menentukan sejauh mana akurasi metode *Naive Bayes Classifier* dalam mengklasifikasikan setimen ulasan pengguna terhadap produk kecantikan.
2. **Pengumpulan Data**
Penelitian ini memanfaatkan data ulasan produk kecantikan jenis *moisturizer* dari *brand* Skintific yang diperoleh melalui *platform* *Female Daily*. Data tersebut dikumpulkan menggunakan teknik *scraping* dengan bantuan ekstensi Instant Data Scraper. Sebanyak 1050 ulasan berhasil diperoleh dan disimpan dalam file *CSV*.
3. **Preprocessing**
Penelitian ini melalui beberapa tahap *preprocessing*, diantaranya:
 - a. **Data Cleaning**
Proses pembersihan data mencakup penghilangan data duplikat, kosong, dan data yang tidak akurat. Jika data tidak dibersihkan, hal ini dapat mengganggu interpretasi hasil visualisasi dan analisis, serta menyebabkan kesalahan dalam pengambilan keputusan [11].

- b. *Case Folding*
Proses ini mengubah seluruh teks menjadi huruf kecil, karena tidak semua teks menggunakan huruf kapital secara konsisten [12].
 - c. *Normalisasi*
Mengoreksi kata-kata yang tidak baku atau salah ejaan menjadi bentuk yang benar, menghapus tanda baca yang tidak perlu, serta menghilangkan spasi berlebih di awal, akhir, atau tengah teks.
 - d. *Tokenisasi*
Proses membagi teks menjadi bagian-bagian yang lebih kecil dengan tujuan untuk menyederhanakan teks.
 - e. *Remove Stopwords*
Menghilangkan kata-kata yang tidak relevan seperti “dan”, “di”, “yang”, dan sebagainya.
 - f. *Stemming*
Proses menganalisis setiap kata yang telah melalui tahap *remove stopwords* untuk mendapatkan bentuk kata dasarnya.
4. *Data Labelling*
Pada tahap ini proses *labelling sentiment* menggunakan Vader Sentiment, Vader Sentiment bekerja dengan baik pada teks dalam bahasa Inggris, karena model ini dilatih untuk menangani sentimen pada kalimat yang menggunakan bahasa sehari-hari dalam bahasa Inggris. Oleh karena itu, untuk mendapatkan analisis yang lebih akurat, teks yang bukan bahasa Inggris perlu diterjemahkan terlebih dahulu ke dalam bahasa Inggris.
 5. *Implementasi Naïve Bayes Classifier*
Tahap ini menerapkan *Naïve Bayes Classifier* untuk mengklasifikasi sentimen ulasan produk kecantikan. Setelah data melalui tahap *preprocessing* dan *labelling*, data dipecah menjadi data latih dan data uji dengan proporsi 80:20 untuk melatih model. Model *Naïve Bayes Classifier* kemudian dilatih pada data latih yang telah dilabeli, dan hasil klasifikasi diuji pada data uji untuk mengevaluasi akurasi prediksi.

C. Hasil dan Pembahasan

1. *Preprocessing*

Penelitian ini melalui beberapa tahapan *preprocessing*, dimulai dengan tahap *data cleaning* yang bertujuan menghapus data duplikat dan data kosong. Dari total 1050 data awal, terdapat 8 data yang terdeteksi duplikat. Setelah proses *data cleaning* dilakukan, jumlah data berkurang menjadi menjadi 1042. Data yang akan digunakan untuk tahap analisis lebih lanjut menjadi lebih bersih dan berkualitas. Tahap kedua yaitu melakukan *case folding* atau mengubah huruf menjadi huruf kecil. Pada gambar 1 berisi data awal sebelum melakukan *preprocessing*, pada Tabel 1 manampilkan data sebelum dan sesudah melakukan *case folding*.

	profile-username	profile-age	review-date	text-content
0	rfistaas	19 - 24	23-Jun-23	di aku gak cocok. mungkin gegara mukaku sensit...
1	Cndycandy	25 - 29	24-Dec-22	Pemah sempet coba yang travel size, oke oke a...
2	zeeklaveya	25 - 29	30-Nov-22	di kulitku yg combi kurang melembabkan dan pak...
3	nishadwii	19 - 24	23-Dec-22	pertama kali tau produk ini soalnya viral bang...
4	ChickenFried20	25 - 29	7-Feb-23	Skintific 5x ceramide ini adalah penyebab kuli...

Gambar 1. Data Asli

Tabel 1. Case Folding

Sebelum	Sesudah
<p>Produknya viral banget aku keracunan dari sosmed. Teksturnya ringan dan mudah meresap, tapi di kulit aku bikin bruntusan sama kusam. wanginya lembut dan packagingnya lucu . setelah pemakaian kulit jadi lembut, banyak promo tapi sayang ga cocok di kulit aku.</p> <p>Banyak banget yang suka sama moisturizer viral ini. Texture wise emang enak banget diaplikasi ke wajah. Gel cream texture, langsung menyerap dan melembapkan. Butuh waktu beberapa saat untuk layer sunscreen setelahnya, karena finishnya basah. Pelembap yang oke tapi ga se wow itu di kulitku. Tapi tetep akan aku habisin 1 jar, walaupun sepertinya tidak akan aku repurchase.</p>	<p>produknya viral banget aku keracunan dari sosmed teksturnya ringan dan mudah meresap tapi di kulit aku bikin bruntusan sama kusam wanginya lembut dan packagingnya lucu setelah pemakaian kulit jadi lembut banyak promo tapi sayang ga cocok di kulit aku</p> <p>banyak banget yang suka sama moisturizer viral ini texture wise emang enak banget diaplikasi ke wajah gel cream texture, langsung menyerap dan melembapkan. butuh waktu beberapa saat untuk layer sunscreen setelahnya, karena finishnya basah pelembap yang oke tapi ga se wow itu di kulitku tapi tetep akan aku habisin 1 jar walaupun sepertinya tidak akan aku repurchase</p>

Langkah selanjutnya adalah normalisasi pada teks dilakukan, yang bertujuan mengubah kata tidak baku menjadi kata yang sesuai dengan kaidah bahasa Indonesia, sehingga makna kata menjadi lebih jelas dan seragam. Contoh beberapa kata baku dan tidak baku disajikan pada Tabel 2.

Tabel 2. Contoh Kata Baku dan Kata Tidak Baku

Kata Tidak Baku	Kata Baku
yg	yang
bgt	sangat
krn	karena
ga, ngga, nggak, gak, gk	tidak
hrs	harus

Tahapan selanjutnya adalah Tokenisasi, yang bertujuan membagi teks menjadi bagian-bagian yang lebih kecil, seperti kata per kata. Proses ini dilakukan untuk menyederhanakan analisis dan membuat model lebih mudah memproses teks dalam bentuk kata-kata individual. Berikut pada Tabel 3 merupakan hasil proses tokenisasi.

Tabel 3. Tokenisasi

Dataset Ulasan Produk	Tokenisasi
produknya viral banget aku keracunan dari	produknya, viral, banget, aku, keracunan, dari,

<p>sosmed teksturnya ringan dan mudah meresap tapi di kulit aku bikin bruntusan sama kusam wanginya lembut dan packagingnya lucu setelah pemakaian kulit jadi lembut banyak promo tapi sayang ga cocok di kulit aku</p>	<p>sosmed, teksturnya, ringan, dan, mudah, meresap, tapi, di, kulit, aku, bikin, bruntusan, sama, kusam, wanginya, lembut, dan, packagingnya, lucu, setelah, pemakaian, kulit, jadi, lembut, banyak, promo, tapi, sayang, ga, cocok, di, kulit, aku,</p>
<p>banyak banget yang suka sama moisturizer viral ini texture wise emang enak banget diaplikasi ke wajah gel cream texture, langsung menyerap dan melembapkan. butuh waktu beberapa saat untuk layer sunscreen setelahnya, karena finishnya basah pelembap yang oke tapi ga se wow itu di kulitku tapi tetep akan aku habisin 1 jar walaupun sepertinya tidak akan aku repurchase</p>	<p>banyak, banget, yang, suka, sama, moisturizer, viral, ini, texture, wise, emang, enak, banget, diaplikasi, ke, wajah, gel, cream, texture, langsung, menyerap, dan, melembapkan, butuh, waktu, beberapa, saat, untuk, layer, sunscreen, setelahnya, karena, finishnya, basah, pelembap, yang, oke, tapi, ga, se, wow, itu, di, kulitku, tapi, tetep, akan, aku, habisin, 1, jar, walaupun, sepertinya, tidak, akan, aku, repurchase,</p>

Kemudian, dilakukan proses *remove stopwords* untuk menghilangkan kata-kata yang kurang penting atau tidak relevan, seperti “dan”, “yang”, atau “dari”. Proses ini membantu model untuk lebih fokus pada kata-kata yang memiliki makna signifikan terhadap sentimen. Hasil dari proses *remove stopwords* disajikan pada Tabel 4.

Tabel 4. Hasil *Remove Stopwords*

Dataset Hasil Remove Stopwords

<p>produk viral banget racun sosmed tekstur ringan mudah resap kulit bikin bruntusan kusam wangi lembut packagingnya lucu pakai kulit lembut promo sayang cocok kulit</p>
<p>bangat suka moisturizer viral texture wise emang enak banget aplikasi wajah gel cream texture langsung serap lembap butuh layer sunscreen telah finishnya basah lembap oke tidak wow kulit tetep habisin 1 jar tidak repurchase</p>

Tahapan terakhir adalah *stemming*, yaitu mengubah kata menjadi bentuk dasar atau kata dasar. Proses ini bertujuan untuk menyederhanakan kata-kata ke bentuk yang lebih umum sehingga makna dari kata tersebut tidak berubah, namun lebih mudah diproses oleh model. Tabel 5 memperlihatkan hasil *stemming*, di mana setiap kata dikembalikan ke bentuk dasarnya.

Tabel 5. Hasil *Stemming*

Dataset Hasil Stemming

<p>produk viral banget racun sosmed tekstur ringan mudah resap kulit bikin beruntusan kusam wangi lembut packagingnya lucu pakai kulit lembut promo sayang cocok kulit</p>
<p>bangat suka moisturizer viral texture wise emang enak banget aplikasi wajah gel cream texture langsung serap lembap butuh layer sunscreen telah finishnya basah lembap oke wow kulit tetep habisin 1 jar repurchase</p>

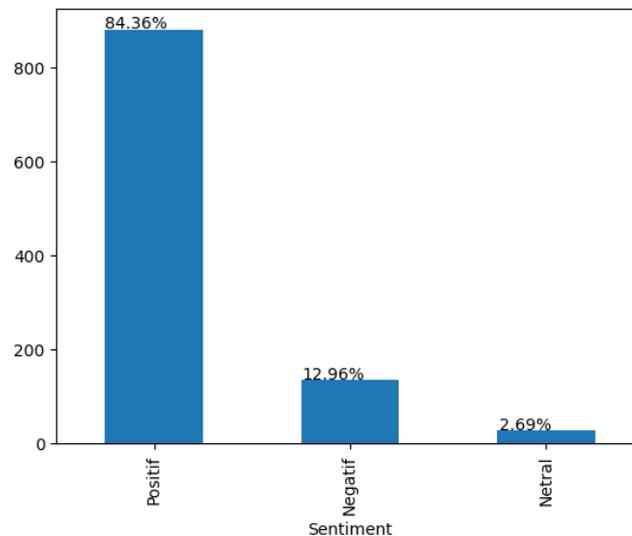
2. Data Labelling

Setelah data melalui tahapan *preprocessing*, langkah berikutnya adalah proses *data labelling*. Data yang sudah bersih, dari 1.050 data awal menjadi 1.042 data, selanjutnya diberi label positif, negatif, atau netral. Untuk mendapatkan hasil akurasi yang akurat, data ulasan ditranslate dari bahasa Indonesia ke bahasa Inggris, karena pada proses pelabelan data yang digunakan adalah kamus Vader

Sentiment. Vader sentiment ini merupakan salah satu dari kamus berbasis lexicon yang menggunakan bahasa Inggris [13], setelah data di *translate* baru akan di beri label. Berikut hasil labelling disajikan pada Tabel 6 dan hasil klasifikasi sentimen divisualisasikan pada Gambar 2.

Tabel 6. *Data Labelling*

Dataset Ulasan Produk	Label
very viral product, social media poison, light texture, easy to absorb into skin, makes rashes dull, soft fragrance, cute packaging, uses soft skin, darling promo, suits skin	Negatif
i really like the viral texture wise moisturizer, it's really delicious, facial application, gel cream texture, absorbs moisture immediately, needs a layer of sunscreen, the finish is wet, moist, okay, wow, my skin still uses up 1 jar, repurchase	Positif



Gambar 2. Klasifikasi sentimen

3. Implementasi *Naïve Bayes Classifier*

Sebelum proses pengujian, diperlukan membagi dataset yang telah diolah menjadi dua bagian, yaitu data latih dan data uji dengan proporsi 80:20. Karena adanya ketidakseimbangan data, diperlukan teknik khusus agar model tidak terlalu fokus pada kelas dominan, yang dapat menyebabkan hasil menjadi bias. Teknik *SMOTE* diterapkan sebagai metode oversampling tanpa mengurangi jumlah data. Setelah proses *SMOTE*, dilakukan pencarian *hyperparameter* menggunakan *GridSearchCV*, yang mengevaluasi akurasi untuk setiap kombinasi parameter dalam *grid*, guna menemukan nilai parameter optimal yang menghasilkan tingkat akurasi tertinggi [14]. Ketika model *Naïve Bayes* dilatih hanya dengan menggunakan *SMOTE* tanpa *GridSearchCV*, akurasi yang diperoleh adalah 79%. Hasil ini lebih rendah dibandingkan dengan akurasi model yang dilatih menggunakan *GridSearchCV*, yang mencapai akurasi 82,77%. Dengan demikian penerapan *GridSearchCV* memiliki peran penting dalam mengoptimalkan performa model

melalui pencarian hyparameter terbaik, sehingga hasil klasifikasi menjadi lebih akurat dan stabil. Tabel 7 dan Tabel 8 menyajikan hasil evaluasi performa model, mencakup akurasi, precision, recall, dan f1-score sebelum dan sesudah penggunaan *GridSearchCV*.

Tabel 7. Sebelum menggunakan *GridSearchCV*

Akurasi	79%
Precision	83%
Recall	79%
F1-score	80%

Tabel 8. Sesudah menggunakan *GridSearchCV*

Akurasi	82,77%
Precision	79%
Recall	83%
F1-score	81%

D. Simpulan

Penelitian analisis sentimen produk kecantikan pada *platform Female Daily* dengan data ulasan sebanyak 1.050 data, setelah dilakukan *preprocessing* menghasilkan data bersih sebanyak 1.042 data dengan klasifikasi sentimen positif sebanyak 84.36%, sentimen negatif sebanyak 12.96%, dan sentimen netral sebanyak 2.69%. Teknik *SMOTE* diterapkan untuk mengatasi ketidakseimbangan data guna mengurangi kecenderungan model pada kelas mayoritas, sementara penggunaan *GridSearchCV* memungkinkan pencarian *hyperparameter* optimal, yang berfungsi untuk memaksimalkan akurasi model. Hasil menunjukkan bahwa akurasi model yang dilatih dengan *GridSearchCV* mencapai 82.77%, lebih tinggi dibandingkan akurasi 79% pada model tanpa penerapan *GridSearchCV*.

E. Ucapan Terima Kasih

Terima kasih disampaikan oleh penulis kepada seluruh pihak yang telah memberikan bantuan selama proses penelitian. Dukungan dari seluruh pihak yang membantu sangat berarti dalam kelancaran dan keberhasilan penelitian ini.

F. Referensi

- [1] L. Fadhillah, "Analisis Media Monitoring terhadap Produk Baru Brand Skintific pada Bulan Maret 2024," *J. Bisnis dan Komun. Digit.*, vol. 1, no. 3, p. 10, 2024, doi: <https://doi.org/10.47134/jbkd.v1i3.2754>.
- [2] D. Lestari, A. Asriani, N. W. Ningsih, W. Rosilawati, and D. H. Amrina, "Pengaruh Kualitas Produk, Fitur Produk dan Desain Produk Terhadap Keputusan Pembelian Produk Kosmetik Dalam Perspektif Manajemen Islam," *REVENUE J. Manaj. Bisnis Islam*, vol. 3, no. 1, pp. 57–80, 2022, doi: <https://dx.doi.org/10.24042/revenue.v3i1.10321>.
- [3] A. W. Prasetyo and R. E. Putra, "Analisis Sentimen Twitter Pada Cryptocurrency Menggunakan Bidirectional Encoder Representations From Transformers Dan Region-Based Convolutional Neural Network," *J. Informatics Comput. Sci.*, pp. 11–21, 2024, doi: <https://ejournal.unesa.ac.id/index.php/jinacs/issue/view/3090>.

- [4] D. M. Azzahra, M. T. Hafid, and S. Alam, "Analisis Sentimen Ulasan Produk Serum Wajah Pada Beauty Brand Somethinc Menggunakan Metode Naïve Bayes Classifier," *JATI (Jurnal Mhs. Tek. Inform.,* vol. 7, no. 3, pp. 1604–1611, 2023, doi: [10.36040/jati.v7i3.6929](https://doi.org/10.36040/jati.v7i3.6929), doi: <https://ejournal.itn.ac.id/index.php/jati/issue/view/291>.
- [5] R. L. Simanjuntak, T. R. Siagian, V. Anggriani, and A. Arnita, "Analisis Sentimen Ulasan Pada Aplikasi E-Commerce Shopee Dengan Menggunakan Algoritma Naïve Bayes," *J. Tek. Mesin, Elektro dan Ilmu Komput.,* vol. 3, no. 3, pp. 23–39, 2023, doi: <https://doi.org/10.55606/teknik.v3i3.2411>.
- [6] G. R. Ramadhan and C. A. Sugianto, "Analisis Sentimen Ulasan Aplikasi Dana di Google Play Store Menggunakan Algoritma Naive Bayes," *JATI (Jurnal Mhs. Tek. Inform.,* vol. 8, no. 5, pp. 9849–9857, 2024, doi: <https://doi.org/10.36040/jati.v8i5.10732>.
- [7] A. Nugroho and Y. Religia, "Analisis Optimasi Algoritma Klasifikasi Naive Bayes menggunakan Genetic Algorithm dan Bagging," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi),* vol. 5, no. 3, pp. 504–510, 2021, doi: <https://doi.org/10.29207/resti.v5i3.3067>.
- [8] T. Astuti and Y. Astuti, "Analisis Sentimen Review Produk Skincare Dengan Naïve Bayes Classifier Berbasis Particle Swarm Optimization (PSO)," *J. Media Inform. Budidarma,* vol. 6, no. 4, pp. 1806–1815, 2022, doi: <http://dx.doi.org/10.30865/mib.v6i4.4119>.
- [9] M. Hamka, N. Alfatari, and D. R. Sari, "Analisis Sentimen Produk Kecantikan Jenis Serum Menggunakan Algoritma Naïve Bayes Classifier," *J. Sist. Komput. dan Inform.,* vol. 4, no. 1, pp. 64–71, 2022, doi: <http://dx.doi.org/10.30865/json.v4i1.4740>.
- [10] Q. A. N. Prakoso, A. Muliawati, and I. N. Isnainiyah, "Analisis Sentimen terhadap Produk Skin Game di Forum Review Female Daily Menggunakan Metode Multinomial Naïve Bayes dan TF-IDF," *Inform. J. Ilmu Komput.,* vol. 18, no. 3, pp. 198–207, 2022, doi: <https://doi.org/10.52958/iftk.v18i3.4679>.
- [11] A. Putri, "Pentingnya data cleaning sebelum visualisasi: Teknik dan tips," *J. Teknol. Pint.,* vol. 4, no. 5, 2024, <http://teknologipintar.org/index.php/teknologipintar/issue/view/34>.
- [12] D. Ananda and R. R. Suryono, "Analisis Sentimen Publik Terhadap Pengungsi Rohingya di Indonesia dengan Metode Support Vector Machine dan Naïve Bayes," *J. MEDIA Inform. BUDIDARMA,* vol. 8, no. 2, pp. 748–757, 2024, doi: <http://dx.doi.org/10.30865/mib.v8i2.7517>.
- [13] N. Habibah, E. Budianita, M. Fikry, and I. Iskandar, "Analisis Sentimen Mengenai Penggunaan E-Wallet Pada Google Play Menggunakan Lexicon Based dan K-Nearest Neighbor," *JURIKOM (Jurnal Ris. Komputer),* vol. 10, no. 1, pp. 192–200, 2023, doi: <http://dx.doi.org/10.30865/jurikom.v10i1.5429>.
- [14] W. Aprilliandhika and F. F. Abdulloh, "Comparison Of K-Nearest Neighbor And Support Vector Machine Algorithm Optimization With Grid Search CV On Stroke Prediction," *J. Tek. Inform.,* vol. 5, no. 4, pp. 991–1000, 2024.