

## **Analisis Sentimen Ulasan Produk Sunscreen Wardah Pada *Marketplace* Shopee Menggunakan Metode *Naïve Bayes***

**<sup>1</sup>Nurhayati Rambe, <sup>2</sup>Syaiful Zuhri Harahap, <sup>3</sup>Ali Akbar Ritonga, <sup>4</sup>Budianto Bangun**

<sup>1,2,4</sup>Sistem Informasi, Fakultas Sains dan Teknologi, Universitas Labuhanbatu

<sup>3</sup>Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas Labuhanbatu

Email : [<sup>1</sup>nrambe155@gmail.com](mailto:nrambe155@gmail.com), [<sup>2</sup>syaifulzuhriharahap@gmail.com](mailto:syaifulzuhriharahap@gmail.com),  
[<sup>3</sup>aliakbarritonga@gmail.com](mailto:aliakbarritonga@gmail.com), [<sup>4</sup>budiantobangun44@gmail.com](mailto:budiantobangun44@gmail.com)

*Corresponding Author* : [nrambe155@gmail.com](mailto:nrambe155@gmail.com)

### ***Abstract***

*The development of the digital world and the popularity of online marketplaces such as Shopee have changed the way consumers interact and review products. Reviews of Wardah sunscreen products, which have an important role in skin health, are one of the most widely found. Understanding the sentiment of these reviews is crucial for manufacturers to improve product quality. Therefore, this study aims to analyze and classify consumer sentiment towards Wardah sunscreen products on Shopee. Using the Naïve Bayes classification method, the reviews will be categorized into positive, negative, and neutral sentiments to get an overall picture of the public perception of the product.*

***Keywords*** : *Sentiment Analysis, Naïve Bayes, Sunscreen Wardah, Shopee.*

### **1. Pendahuluan**

Perkembangan pesat dunia digital telah melahirkan pergeseran paradigma dalam perilaku konsumen, terutama dengan popularitas marketplace online seperti Shopee di Indonesia. Platform ini tidak hanya berfungsi sebagai etalase virtual, tetapi juga sebagai ruang publik di mana konsumen dapat secara aktif memberikan umpan balik dalam bentuk ulasan produk. Ulasan ini telah menjadi aset digital yang sangat berpengaruh, karena menjadi acuan utama bagi calon pembeli dalam membuat keputusan. Di tengah persaingan pasar yang ketat, kemampuan perusahaan untuk memahami dan merespons sentimen konsumen yang terkandung dalam ulasan menjadi faktor penentu keberhasilan. Salah satu kategori produk yang mendapat perhatian signifikan di platform ini adalah sunscreen atau tabir surya. Mengingat peran esensialnya dalam melindungi kulit dari dampak buruk paparan sinar matahari, permintaan dan diskusi seputar produk ini terus meningkat. Wardah, sebagai merek lokal yang dominan, menjadi subjek studi kasus yang relevan karena popularitasnya yang tinggi. Ribuan ulasan yang ditinggalkan konsumen pada produk sunscreen Wardah di Shopee merupakan kumpulan data yang kaya akan opini, persepsi, dan sentimen. Namun, volume data tekstual yang tidak terstruktur ini menghadirkan tantangan dalam ekstraksi informasi yang bermakna. Untuk mengatasi tantangan tersebut, penelitian ini mengusulkan penerapan analisis sentimen menggunakan metode *Naïve Bayes Classifier*. Analisis sentimen adalah

teknik komputasi linguistik untuk mengidentifikasi dan mengekstrak opini secara otomatis dari sebuah teks. Metode Naïve Bayes dipilih karena keandalannya dalam menangani masalah klasifikasi teks, efisiensi komputasi, dan kinerjanya yang baik bahkan dengan dataset yang tidak terlalu besar. Penelitian ini secara spesifik bertujuan untuk mengklasifikasikan sentimen ulasan konsumen terhadap produk sunscreen Wardah di Shopee ke dalam tiga kategori: positif, negatif, dan netral. Dengan demikian, penelitian ini diharapkan dapat memberikan wawasan empiris mengenai penerimaan pasar terhadap produk tersebut dan menunjukkan bagaimana metode komputasional dapat dimanfaatkan untuk mengubah data ulasan kualitatif menjadi metrik kuantitatif yang dapat ditindaklanjuti oleh perusahaan.

## **2. Landasan Teori**

### **Data Mining**

Data mining adalah proses menemukan pola atau pengetahuan tersembunyi dalam data. Dalam konteks ulasan produk, data mining digunakan untuk mengeksplorasi dataset besar seperti ulasan konsumen, dengan tujuan untuk mengekstrak informasi yang dapat digunakan untuk meningkatkan produk atau layanan. Data mining melibatkan tiga komponen utama: eksplorasi data, pemodelan prediktif, dan evaluasi hasil. Penggunaan teknik data mining dalam analisis ulasan produk telah terbukti efektif dalam mengidentifikasi preferensi konsumen serta pola perilaku yang dapat digunakan untuk pengembangan produk (Han et al., 2012).

### ***Knowledge Discovery in Database***

KDD adalah proses yang lebih luas dari data mining yang mencakup identifikasi pola-pola baru dalam data yang sebelumnya tidak diketahui. Proses ini meliputi beberapa tahap, seperti seleksi data, pra-pemrosesan, transformasi data, dan evaluasi pola. KDD bertujuan untuk menemukan pengetahuan yang dapat digunakan untuk keputusan strategis. *Knowledge Discovery in Database* adalah pendekatan yang sangat penting dalam analisis data besar dan memiliki aplikasi luas di berbagai bidang, termasuk analisis sentimen produk (Fayyad et al., 1996).

### **Analisis Sentimen**

Analisis sentimen adalah teknik untuk mengidentifikasi, mengukur, dan mengklasifikasikan perasaan atau opini yang terkandung dalam teks. Dalam ulasan produk, analisis sentimen dapat digunakan untuk mengetahui apakah konsumen memiliki sentimen positif, negatif, atau netral terhadap produk yang mereka beli. Analisis sentimen memainkan peran penting dalam memahami persepsi pelanggan, terutama di platform e-commerce. Dengan memanfaatkan metode seperti Naïve Bayes, kita dapat mengotomatisasi analisis sentimen dalam ulasan produk untuk memberikan wawasan yang lebih cepat dan lebih akurat (Pang & Lee, 2008).

### **Produk Sunscreen Wardah**

Wardah adalah salah satu merek kosmetik terkemuka di Indonesia yang memiliki berbagai produk perawatan kulit, termasuk sunscreen. Sunscreen Wardah dipercaya memiliki manfaat dalam melindungi kulit dari paparan sinar UV yang berbahaya. Bahwa konsumen di Indonesia semakin sadar akan pentingnya perlindungan kulit dari sinar matahari, sehingga permintaan terhadap produk sunscreen terus meningkat. Evaluasi melalui analisis sentimen dapat membantu Wardah memahami penerimaan konsumen terhadap produk sunscreen mereka (Efrina & Lusiana, 2025).

### **Marketplace Shopee**

Shopee merupakan salah satu platform e-commerce yang paling berkembang di Indonesia. Marketplace ini menyediakan berbagai produk, termasuk produk kecantikan seperti sunscreen Wardah. Shopee menjadi *platform* yang dominan dalam jual beli produk kecantikan karena kemudahan akses dan berbagai promo yang ditawarkan. Ulasan konsumen di Shopee sering kali menjadi bahan pertimbangan penting bagi calon pembeli (Aghniyatussalma, 2022).

### **Metode Naïve Bayes**

*Metode Naïve Bayes* adalah salah satu teknik dalam data mining yang sering digunakan untuk analisis teks, termasuk analisis sentimen. Metode ini berdasarkan pada teorema Bayes dan mengasumsikan independensi antar fitur yang ada. Naïve Bayes efektif untuk menangani masalah klasifikasi, termasuk dalam kasus analisis sentimen terhadap produk di marketplace. Teknik ini telah terbukti dapat memberikan hasil yang baik meskipun dengan asumsi kesederhanaannya (Parsons, 2010).

## **3. Metode Penelitian**

### **Desain Penelitian**

Penelitian ini menggunakan pendekatan kuantitatif dengan metode analisis sentimen untuk mengeksplorasi opini konsumen mengenai produk Sunscreen Wardah di platform e-commerce Shopee. Desain penelitian menggabungkan teknik data mining dan analisis sentimen, bertujuan untuk mengidentifikasi pola atau sentimen yang terkandung dalam ulasan produk.

### **Populasi dan Sampel**

Populasi penelitian ini mencakup seluruh ulasan konsumen terhadap produk Sunscreen Wardah yang ada di Shopee. Sampel yang digunakan adalah ulasan yang diposting dalam enam bulan terakhir, diambil secara acak dengan kriteria validitas tertentu, seperti ulasan yang mencakup teks dan rating yang jelas. Jumlah sampel disesuaikan berdasarkan analisis awal untuk memastikan representativitas terhadap populasi ulasan produk.

### **Instrumen Penelitian**

Instrumen utama dalam penelitian ini adalah ulasan produk Sunscreen Wardah yang tersedia di platform Shopee. Ulasan ini akan dianalisis untuk mengevaluasi sentimen konsumen.

## Alat Analisis Sentimen

Analisis sentimen dilakukan menggunakan perangkat lunak *RapidMiner* dengan algoritma Naïve Bayes untuk mengklasifikasikan sentimen ulasan menjadi tiga kategori: positif, negatif, atau netral.

## Variabel Penelitian

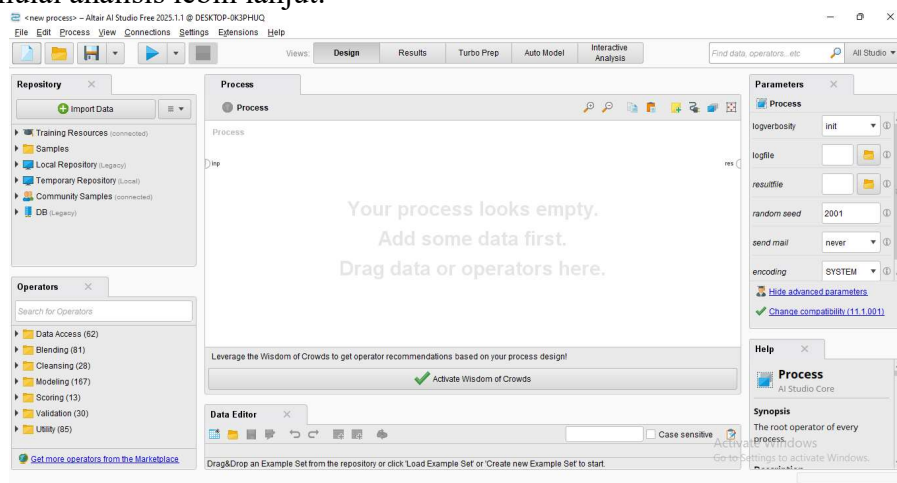
- Variabel Independen (*Input*) Isi teks ulasan yang diberikan oleh konsumen pada produk Sunscreen Wardah.
- Variabel Dependen (*Output*) Klasifikasi sentimen dari ulasan konsumen, yang dianalisis menggunakan metode Naïve Bayes (positif, negatif, netral).

## Metode Pengumpulan Data

Metode pengumpulan data dilakukan dengan teknik web scraping atau API Shopee untuk mengakses ulasan produk Sunscreen Wardah. Data yang dikumpulkan meliputi teks ulasan dan rating yang diberikan oleh konsumen.

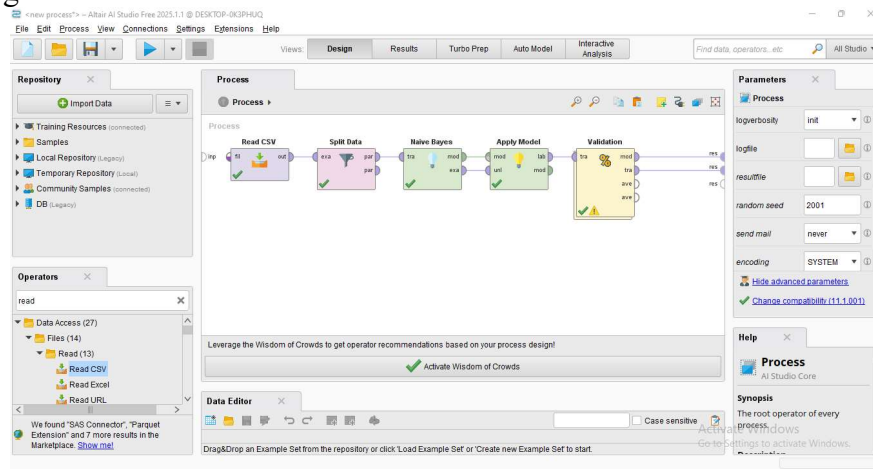
## 4. Hasil dan Pembahasan Pembahasan

Tampilan awal dari Altair AI Studio, di mana pengguna mulai dengan proyek kosong. Pada tampilan ini, pengguna dapat melihat area kerja utama dengan panel Repository di sebelah kiri, yang menyediakan akses ke berbagai sumber daya data, seperti Training Resources dan Samples. Di bagian tengah, area Process menunjukkan tampilan kosong, yang menandakan bahwa proses analisis atau pemodelan belum dimulai. Pengguna diberi instruksi untuk menambahkan data atau operator terlebih dahulu untuk memulai alur kerja. Pada bagian kanan, terdapat panel Parameters yang memungkinkan pengguna untuk mengonfigurasi berbagai pengaturan seperti logverbosity dan random seed. Tampilan ini menandakan bahwa pengguna perlu mengimpor data atau memilih operator yang relevan untuk memulai analisis lebih lanjut.



**Gambar 1. Tampilan Awal Proses di Altair AI Studio**

Alur proses analisis sentimen yang dilakukan menggunakan Altair AI Studio. Proses dimulai dengan operator Read CSV, yang digunakan untuk mengimpor data ulasan produk dalam format CSV. Selanjutnya, data tersebut dibagi menjadi dua subset menggunakan operator Split Data, yang memisahkan data menjadi bagian untuk pelatihan dan pengujian. Pada langkah berikutnya, model *Naïve Bayes* diterapkan pada data latih menggunakan operator Naïve Bayes, untuk membangun model klasifikasi sentimen berdasarkan ulasan yang ada. Setelah model dilatih, operator *Apply Model* digunakan untuk menerapkan model tersebut pada data uji guna memprediksi sentimen dari ulasan baru. Terakhir, operator Validation digunakan untuk mengevaluasi performa model dengan menghasilkan metrik evaluasi, seperti akurasi dan presisi, untuk mengukur kualitas model dalam mengklasifikasikan sentimen ulasan. Proses ini memberikan gambaran sistematis dari tahapan analisis sentimen yang dilakukan untuk memahami persepsi konsumen terhadap produk yang dianalisis.



**Gambar 2. Proses Analisis Sentimen Menggunakan Naïve Bayes di Altair AI Studio**

Hasil prediksi sentimen untuk data ulasan produk yang telah dianalisis menggunakan model Naïve Bayes di Altair AI Studio. Setiap baris mewakili satu ulasan yang telah diklasifikasikan ke dalam kategori sentimen: Positif, Negatif, atau Netral. Kolom "prediction(S)" menunjukkan prediksi model untuk masing-masing ulasan, sementara kolom "confidence" menggambarkan tingkat keyakinan model terhadap klasifikasi tersebut, dengan nilai yang lebih tinggi menunjukkan kepercayaan yang lebih besar dalam prediksi yang diberikan. Sebagai contoh, pada baris pertama, model memprediksi sentimen sebagai Positif dengan tingkat kepercayaan sebesar 0.962, yang menunjukkan keyakinan yang tinggi terhadap prediksi tersebut. Di sisi lain, beberapa ulasan yang diprediksi sebagai Negatif memiliki tingkat kepercayaan lebih rendah, misalnya 0.020, yang mencerminkan ketidakpastian dalam prediksi tersebut. Kolom "Ulasan" berisi teks asli dari ulasan yang digunakan untuk klasifikasi. Hasil ini memberikan gambaran tentang bagaimana model dapat memprediksi sentimen berdasarkan teks ulasan yang diberikan.

Row No.	Sentimen	prediction(Sentimen)	confidence(Positif)	confidence(Negatif)	confidence(Neutral)	Ulasan
1	Positif	Positif	0.962	0.020	0.018	Produk ini sa...
2	Negatif	Negatif	0.020	0.961	0.018	Sunscreen in...
3	Netral	Netral	0.022	0.022	0.957	Cukup baik. B...
4	Positif	Positif	0.962	0.020	0.018	Produk ini sa...
5	Negatif	Negatif	0.020	0.961	0.018	Sayang, suns...
6	Netral	Netral	0.022	0.022	0.957	Sunscreen in...
7	Negatif	Negatif	0.020	0.961	0.018	Produk ini bd...
8	Negatif	Negatif	0.020	0.961	0.018	Rasanya terf...
9	Netral	Netral	0.022	0.022	0.957	Produk ini cu...
10	Positif	Positif	0.962	0.020	0.018	Sangat efesi...
11	Negatif	Negatif	0.020	0.961	0.018	Setelah men...
12	Positif	Positif	0.962	0.020	0.018	Sunscreen in...
13	Negatif	Negatif	0.020	0.961	0.018	Saya merasa...
14	Netral	Netral	0.022	0.022	0.957	Produk ini bd...
15	Positif	Positif	0.962	0.020	0.018	Beahju dicaka...

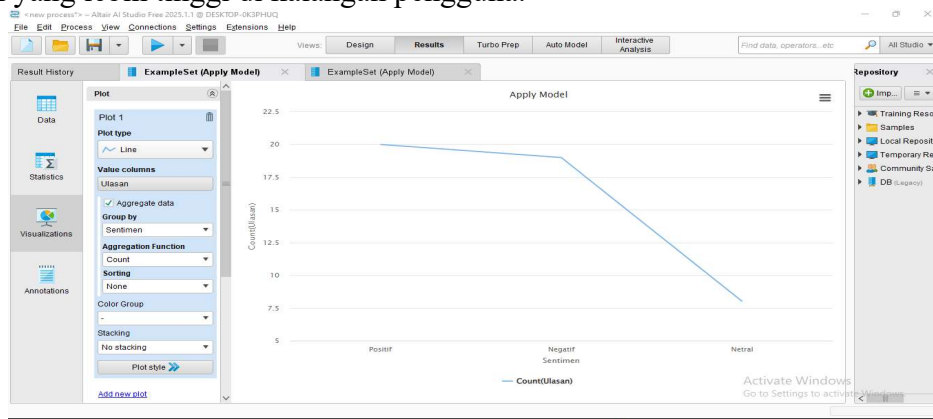
**Gambar 3. Hasil Prediksi Sentimen Menggunakan Model Naïve Bayes di Altair AI Studio**

Statistik hasil prediksi sentimen yang diperoleh setelah penerapan model Naïve Bayes pada data ulasan produk. Tabel ini mencakup beberapa atribut penting yang menggambarkan hasil prediksi dan tingkat kepercayaan model terhadap klasifikasi sentimen. Kolom Sentimen menunjukkan nilai sebenarnya dari sentimen ulasan, yang terbagi menjadi Positif, Negatif, dan Netral. Kolom *Prediction*(Sentimen) memuat hasil prediksi sentimen yang dilakukan oleh model, dengan hasil yang serupa: Positif, Negatif, atau Netral. Selanjutnya, kolom *confidence*(Positif), *confidence*(Negatif), dan *confidence*(Netral) memberikan tingkat keyakinan model terhadap masing-masing kategori sentimen, dengan nilai yang bervariasi dari 0.018 hingga 0.962, menggambarkan tingkat kepercayaan model dalam klasifikasi sentimen tersebut. Kolom Ulasan menampilkan teks asli dari ulasan produk yang digunakan dalam analisis. Statistik tambahan yang ditampilkan, seperti min, max, dan average, memberikan gambaran lebih mendalam mengenai distribusi tingkat kepercayaan yang dihasilkan oleh model pada seluruh dataset yang diuji. Hasil ini menunjukkan efektivitas model dalam mengklasifikasikan ulasan berdasarkan sentimen dengan tingkat kepercayaan yang bervariasi.

Name	Type	Missing	Statistics	Filter (6 / 6 attributes)
Label Sentimen	Polynomial	0	Least: Netral (6) Most: Positif (20)	Values: Positif (20), Negatif (19), ... [1 more]
Prediction prediction(Sentimen)	Polynomial	0	Least: Netral (6) Most: Positif (20)	Values: Positif (20), Negatif (19), ... [1 more]
Confidence_Positif confidence(Positif)	Real	0	Min: 0.020 Max: 0.962	Average: 0.421
Confidence_Negatif confidence(Negatif)	Real	0	Min: 0.020 Max: 0.961	Average: 0.401
Confidence_Neutral confidence(Neutral)	Real	0	Min: 0.018 Max: 0.957	Average: 0.178
Ulasan	Polynomial	0	Least: Warna ku [...] Inif (0) Most: Begitu d [...] dungi (1)	Values: Begitu d [...] rindungi (1), Cukup ba [...]

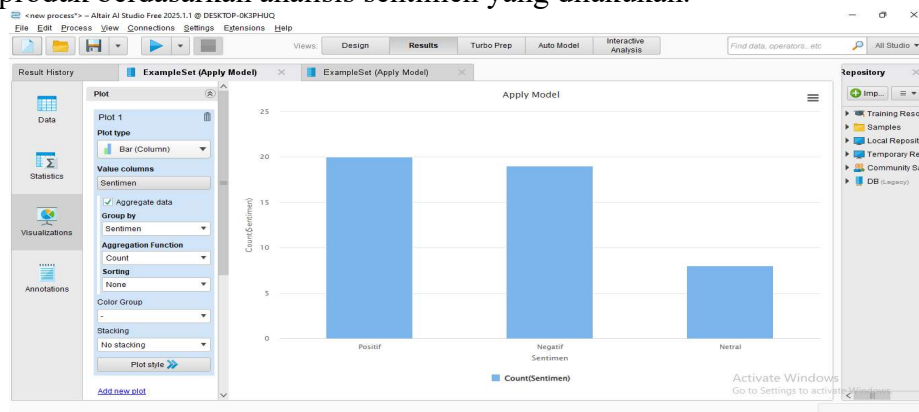
**Gambar 4. Statistik Hasil Prediksi Sentimen Menggunakan Model Naïve Bayes**

Visualisasi distribusi sentimen dari ulasan produk yang telah diproses menggunakan model analisis sentimen. Grafik ini menggunakan sumbu horizontal untuk menunjukkan tiga kategori sentimen: Positif, Negatif, dan Netral. Sumbu vertikal menggambarkan jumlah ulasan yang termasuk dalam masing-masing kategori, yang dihitung menggunakan fungsi agregasi Count. Dalam visualisasi ini, data diatur berdasarkan kategori sentimen, dengan Positif menunjukkan jumlah ulasan terbanyak, diikuti oleh Negatif dan Netral. Grafik ini memberikan gambaran umum tentang persepsi konsumen terhadap produk, dengan lebih banyak ulasan yang cenderung masuk dalam kategori Positif, yang menunjukkan tingkat kepuasan yang lebih tinggi di kalangan pengguna.



**Gambar 5. Visualisasi Distribusi Sentimen Ulasan Produk**

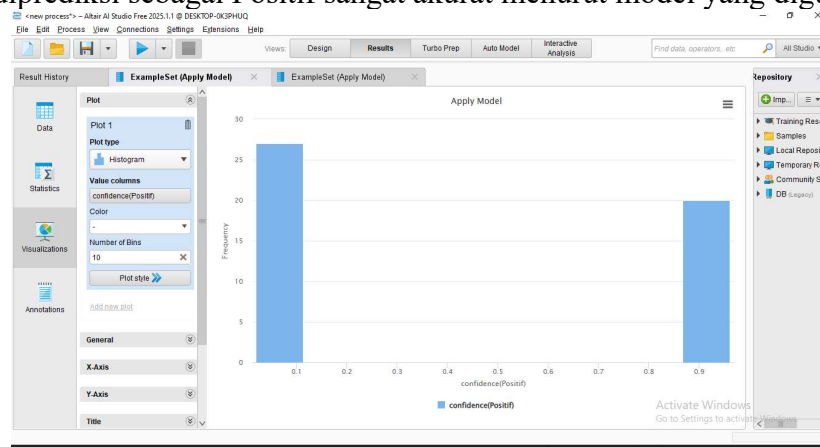
Visualisasi distribusi sentimen dari ulasan produk yang telah dianalisis menggunakan model *Naïve Bayes*. Grafik ini menggunakan grafik batang untuk menggambarkan jumlah ulasan yang terklasifikasi dalam tiga kategori sentimen: Positif, Negatif, dan Netral. Sumbu vertikal menunjukkan jumlah ulasan pada setiap kategori, sedangkan sumbu horizontal menunjukkan tipe sentimen. Berdasarkan visualisasi ini, dapat dilihat bahwa jumlah ulasan dengan sentimen Positif lebih banyak dibandingkan dengan Negatif dan Netral. Hal ini menunjukkan bahwa mayoritas konsumen memberikan ulasan positif terhadap produk tersebut. Grafik ini memberikan gambaran yang jelas tentang persepsi umum konsumen terhadap produk berdasarkan analisis sentimen yang dilakukan.



**Gambar 6. Visualisasi Distribusi Sentimen Ulasan Produk Menggunakan Grafik Batang**



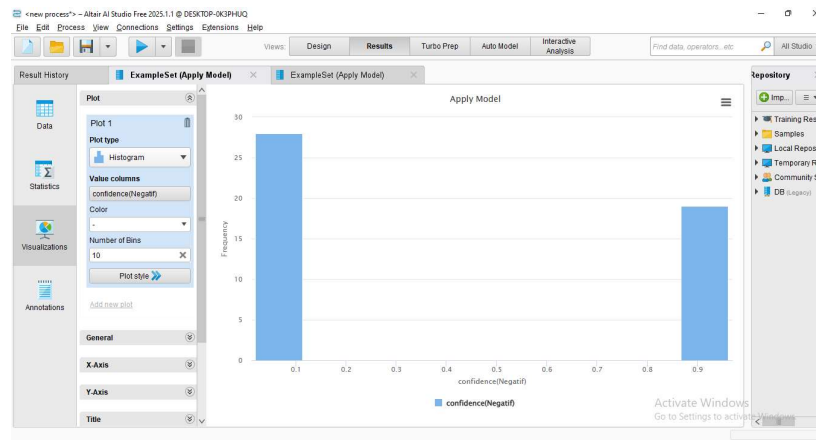
Histogram yang menggambarkan distribusi tingkat kepercayaan model terhadap prediksi sentimen Positif. Sumbu horizontal (X-axis) mewakili nilai tingkat kepercayaan (*confidence*(Positif)), yang berkisar antara 0 dan 1, sementara sumbu vertikal (Y-axis) menunjukkan frekuensi (jumlah) prediksi pada setiap rentang nilai kepercayaan. Grafik ini menggunakan 10 bins untuk membagi rentang nilai kepercayaan menjadi 10 interval, dengan setiap bar mewakili jumlah ulasan yang jatuh dalam interval tertentu. Dapat dilihat bahwa sebagian besar ulasan memiliki tingkat kepercayaan yang sangat tinggi (mendekati 1), yang menunjukkan bahwa model dengan yakin memprediksi ulasan tersebut sebagai Positif. Ini menunjukkan bahwa model memiliki performa yang baik dalam mengklasifikasikan sentimen positif dengan tingkat kepercayaan yang tinggi, mengindikasikan bahwa mayoritas ulasan yang diprediksi sebagai Positif sangat akurat menurut model yang digunakan.



**Gambar 7. Distribusi Tingkat Kepercayaan Model terhadap Sentimen Positif**

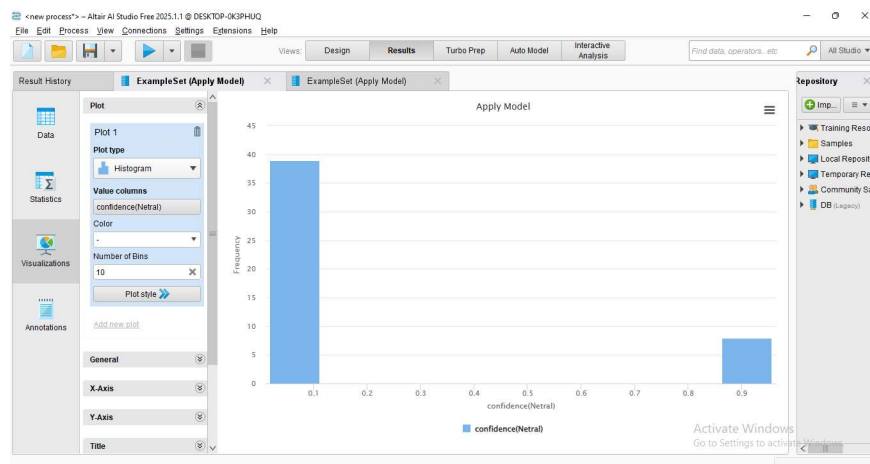
Histogram yang menggambarkan distribusi tingkat kepercayaan model terhadap prediksi sentimen Negatif. Sumbu horizontal (X-axis) mewakili nilai tingkat kepercayaan (*confidence*(Negatif)), yang berkisar antara 0 dan 1, sementara sumbu vertikal (Y-axis) menunjukkan frekuensi atau jumlah ulasan yang terklasifikasi dalam interval nilai tertentu. Grafik ini menggunakan 10 bins untuk membagi rentang nilai kepercayaan menjadi 10 interval, dengan setiap bar mewakili jumlah ulasan yang memiliki tingkat kepercayaan dalam interval tersebut. Dapat dilihat bahwa sebagian besar ulasan yang diprediksi sebagai Negatif memiliki tingkat kepercayaan yang sangat tinggi, mendekati 1, yang menunjukkan bahwa model dengan yakin memprediksi ulasan tersebut sebagai Negatif. Hal ini menunjukkan bahwa model memiliki performa yang baik dalam mengidentifikasi sentimen negatif dengan tingkat keyakinan yang tinggi, mengindikasikan bahwa mayoritas ulasan yang diprediksi sebagai Negatif sangat akurat menurut model yang digunakan.





**Gambar 8. Distribusi Tingkat Kepercayaan Model terhadap Sentimen Negatif**

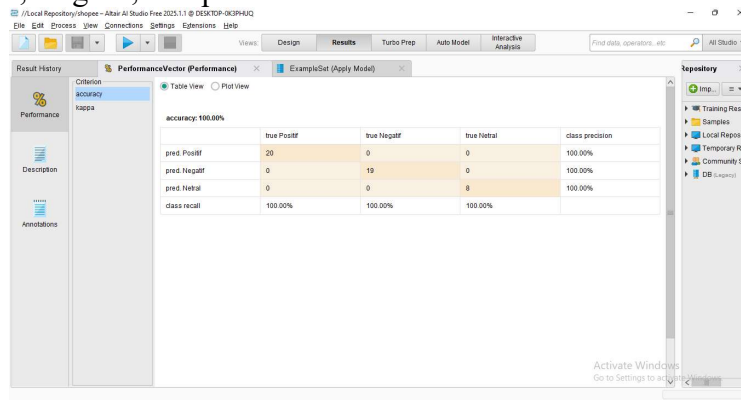
Histogram yang menggambarkan distribusi tingkat kepercayaan model terhadap prediksi sentimen Netral. Sumbu horizontal (X-axis) mewakili tingkat kepercayaan ( $\text{confidence}(\text{Netral})$ ), dengan rentang nilai antara 0 dan 1, sementara sumbu vertikal (Y-axis) menunjukkan frekuensi atau jumlah ulasan yang terklasifikasi dalam interval nilai tertentu. Grafik ini menggunakan 10 bins untuk membagi rentang nilai kepercayaan menjadi 10 interval, dengan setiap bar mewakili jumlah ulasan dalam interval tersebut. Sebagian besar ulasan yang diprediksi sebagai Netral memiliki tingkat kepercayaan yang tinggi, mendekati 1, yang menunjukkan bahwa model dengan yakin mengklasifikasikan ulasan tersebut sebagai Netral. Hal ini menunjukkan bahwa model cukup efektif dalam mengidentifikasi sentimen netral dengan tingkat keyakinan yang konsisten tinggi pada sebagian besar data yang dianalisis.



**Gambar 9. Distribusi Tingkat Kepercayaan Model terhadap Sentimen Netral**

Hasil evaluasi kinerja model Naïve Bayes dalam mengklasifikasikan sentimen ulasan produk menggunakan confusion matrix yang melibatkan metrik accuracy, precision, dan recall untuk setiap kategori sentimen (Positif, Negatif, Netral). Berdasarkan hasil tersebut, model menunjukkan accuracy sebesar 100%, yang menunjukkan bahwa semua prediksi sesuai dengan nilai *true class*. Untuk *class recall*, yang mencerminkan kemampuan model untuk mengidentifikasi setiap kelas dengan benar, hasilnya juga mencapai 100% pada setiap

kategori sentimen: Positif, Negatif, dan Netral. Kolom class precision menunjukkan bahwa model memiliki precision yang sangat tinggi untuk semua kategori, yakni 100%. Ini menunjukkan bahwa model sangat akurat dalam mengklasifikasikan ulasan ke dalam kategori yang benar, tanpa kesalahan dalam setiap kelas. Secara keseluruhan, hasil evaluasi ini mengindikasikan bahwa model Naïve Bayes memiliki performa yang sangat baik dalam klasifikasi sentimen pada dataset yang digunakan, dengan tidak ada kesalahan dalam prediksi baik untuk Positif, Negatif, maupun Netral.



	true Positif	true Negatif	true Netral	class precision
pred. Positif	20	0	0	100.00%
pred. Negatif	0	19	0	100.00%
pred. Netral	0	0	8	100.00%
class recall	100.00%	100.00%	100.00%	

**Gambar 10. Metrik Kinerja Model pada Analisis Sentimen**

Pengujian analisis sentimen menggunakan aplikasi RapidMiner berhasil menunjukkan efektivitas dalam mengklasifikasikan ulasan produk berdasarkan sentimen konsumen (positif, negatif, netral). Proses dimulai dengan pengimporan data ulasan menggunakan operator Read CSV, yang kemudian diproses dan dibersihkan menggunakan teknik pengolahan data teks. Setelah data siap, model Naïve Bayes diterapkan untuk mengklasifikasikan sentimen ulasan berdasarkan kata-kata yang terkandung dalam teks ulasan. Hasil dari model prediksi menunjukkan bahwa aplikasi RapidMiner dapat mengklasifikasikan ulasan dengan tingkat kepercayaan yang tinggi, terutama pada kategori Positif dan Negatif, dimana sebagian besar ulasan mendapatkan prediksi dengan tingkat keyakinan mendekati 1. Hasil visualisasi dan distribusi tingkat kepercayaan juga menunjukkan bahwa model dapat memprediksi dengan akurat, namun ada sebagian kecil ulasan yang memiliki tingkat kepercayaan yang rendah, terutama pada kategori Netral. Secara keseluruhan, penggunaan RapidMiner dalam analisis sentimen ini terbukti efisien dan efektif, dengan kemampuan untuk memberikan gambaran yang jelas mengenai persepsi konsumen terhadap produk. Aplikasi ini memberikan hasil yang dapat diandalkan dalam mengklasifikasikan sentimen, serta menyediakan metrik evaluasi yang berguna untuk menilai kinerja model. Hasil analisis ini dapat digunakan untuk memberikan wawasan yang lebih mendalam dalam pengambilan keputusan bisnis dan strategi pemasaran.

## 5. Kesimpulan dan Saran

### Kesimpulan

1. Pengujian menggunakan aplikasi RapidMiner berhasil menunjukkan bahwa model Naïve Bayes dapat secara efektif mengklasifikasikan ulasan produk menjadi tiga kategori sentimen: positif, negatif, dan netral. Model ini menunjukkan akurasi yang

tinggi, terutama dalam mengklasifikasikan sentimen positif dan negatif, dengan tingkat kepercayaan yang konsisten tinggi. Hasil ini menunjukkan potensi RapidMiner dalam melakukan analisis sentimen secara otomatis dan efisien.

2. Tingkat kepercayaan yang dihasilkan oleh model, khususnya dalam prediksi sentimen positif dan negatif, menunjukkan hasil yang sangat meyakinkan. Namun, tingkat kepercayaan model pada kategori sentimen netral cenderung lebih rendah, menunjukkan bahwa model masih membutuhkan penyempurnaan untuk lebih akurat dalam mengklasifikasikan sentimen netral.
3. Proses pembersihan data yang baik sebelum pemodelan sangat berpengaruh terhadap hasil analisis sentimen. Penggunaan teknik pembersihan data seperti penghapusan stop words, tokenisasi, dan normalisasi teks membantu meningkatkan akurasi prediksi model.

### Saran

1. Untuk meningkatkan akurasi model, disarankan untuk memperbaiki kualitas data yang digunakan dalam pelatihan, terutama dengan menambahkan lebih banyak data ulasan yang mencakup berbagai sentimen, termasuk sentimen netral. Ini akan membantu model untuk lebih memahami variasi dalam teks ulasan dan meningkatkan prediksi pada kategori netral.
2. Meskipun model Naïve Bayes terbukti efektif, disarankan untuk mengeksplorasi model pembelajaran mesin lainnya, seperti Support Vector Machine (SVM) atau Random Forest, untuk membandingkan performa dalam klasifikasi sentimen. Pendekatan ini dapat memberikan wawasan tambahan dan potensi untuk meningkatkan hasil prediksi.
3. Untuk memperoleh gambaran yang lebih komprehensif tentang kinerja model, disarankan untuk melakukan evaluasi model menggunakan metrik tambahan seperti *precision*, *recall*, dan *F1-score*, selain akurasi. Evaluasi ini dapat memberikan pemahaman yang lebih baik tentang keandalan model dalam mengklasifikasikan setiap kategori sentimen secara lebih mendalam.

### 6. Daftar Pustaka

- Aghniyatussalma, A. (2022). FAKTOR-FAKTOR YANG MEMPENGARUHI KEPUTUSAN PEMBELIAN PRODUK KECANTIKAN MELALUI E-COMMERCE SHOPEE. *Blantika: Multidisciplinary Journal*, 1(1), 53–60. <https://doi.org/10.57096/blantika.v1i1.5>
- Alghifari, F., & Juardi, D. (2021). Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes. *Jurnal Ilmiah Informatika*, 9(02), 75–81. <https://doi.org/10.33884/jif.v9i02.3755>
- Efrina, H., & Lusiana, M. (2025). Pengaruh Penggunaan Masker Singkong Kuning terhadap Perawatan Kulit Wajah Kering, Peningkatan Kelembapan, dan Kesehatan Kulit. *Mutiara: Jurnal Penelitian Dan Karya Ilmiah*, 3(4), 252–262. <https://doi.org/10.59059/mutiara.v3i4.2707>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting

- useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27–34. <https://doi.org/10.1145/240455.240464>
- Han, J., Kamber, M., & Pei, J. (2012). Advanced Pattern Mining. In *Data Mining* (pp. 279–325). Elsevier. <https://doi.org/10.1016/b978-0-12-381479-1.00007-1>
- Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2), 1–135. <https://doi.org/10.1561/1500000011>
- Parsons, S. (2010). Introduction to Machine Learning, Second Edition by Ethem Alpaydin, MIT Press, 584 pp., \$55.00. ISBN 978-0-262-01243-0. *The Knowledge Engineering Review*, 25(3), 353. <https://doi.org/10.1017/s0269888910000056>
- Rumah, P., Royal, S., Hidayah, A., Dulisep, D., & Angga, B. (2024). Implementasi Algoritma K - Means Menggunakan RapidMiner untuk Klasterisasi Data Obat. 7(2), 200–211.
- Susilo, J., & Syahra, Y. (2020). Implementasi Data Mining Untuk Menganalisa Pola Penjualan Menu Makanan Berdasarkan Permintaan Konsumen DI Restoran Wakaka Center Point Menggunakan Algoritma Apriori. *Cyber Tech*, 1–14.
- Aghniyatussalma, A. (2022). FAKTOR-FAKTOR YANG MEMPENGARUHI KEPUTUSAN PEMBELIAN PRODUK KECANTIKAN MELALUI E-COMMERCE SHOPEE. *Blantika: Multidisciplinary Journal*, 1(1), 53–60. <https://doi.org/10.57096/blantika.v1i1.5>
- Alghifari, F., & Juardi, D. (2021). Penerapan Data Mining Pada Penjualan Makanan Dan Minuman Menggunakan Metode Algoritma Naïve Bayes. *Jurnal Ilmiah Informatika*, 9(02), 75–81. <https://doi.org/10.33884/jif.v9i02.3755>
- Efrina, H., & Lusiana, M. (2025). Pengaruh Penggunaan Masker Singkong Kuning terhadap Perawatan Kulit Wajah Kering, Peningkatan Kelembapan, dan Kesehatan Kulit. *Mutiara: Jurnal Penelitian Dan Karya Ilmiah*, 3(4), 252–262. <https://doi.org/10.59059/mutiara.v3i4.2707>
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27–34. <https://doi.org/10.1145/240455.240464>
- Han, J., Kamber, M., & Pei, J. (2012). Advanced Pattern Mining. In *Data Mining* (pp. 279–325). Elsevier. <https://doi.org/10.1016/b978-0-12-381479-1.00007-1>
- Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2), 1–135. <https://doi.org/10.1561/1500000011>
- Parsons, S. (2010). Introduction to Machine Learning, Second Edition by Ethem Alpaydin, MIT Press, 584 pp., \$55.00. ISBN 978-0-262-01243-0. *The Knowledge Engineering Review*, 25(3), 353. <https://doi.org/10.1017/s0269888910000056>
- Rumah, P., Royal, S., Hidayah, A., Dulisep, D., & Angga, B. (2024). Implementasi Algoritma K - Means Menggunakan RapidMiner untuk Klasterisasi Data Obat. 7(2), 200–211.
- Susilo, J., & Syahra, Y. (2020). Implementasi Data Mining Untuk Menganalisa Pola Penjualan Menu Makanan Berdasarkan Permintaan Konsumen DI Restoran Wakaka Center Point Menggunakan Algoritma Apriori. *Cyber Tech*, 1–14.