

Analisis Sentimen Ulasan Produk Masker Menggunakan Algoritma Naive Bayes

Fareza Muhammad^{1*}, Devi Yunita²

^{1,2}Ilmu Komputer, Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia

Email: ^{1*}fareza16@gmail.com, ²yunitadevi54@yahoo.co.id

(* : coresponding author)

Abstrak - Penelitian ini menganalisis sentimen ulasan produk masker menggunakan algoritma Naive Bayes. Dataset yang digunakan terdiri dari 99 ulasan produk masker dari platform e-commerce, dengan klasifikasi sentimen positif dan negatif. Metodologi yang digunakan melibatkan preprocessing data, pemodelan menggunakan algoritma Naive Bayes, dan evaluasi performa model. Hasil analisis menunjukkan bahwa algoritma Naive Bayes mampu mengklasifikasikan sentimen ulasan dengan tingkat akurasi 83.67%. Penelitian ini memberikan wawasan berharga bagi pelaku e-commerce dalam memahami persepsi pelanggan terhadap produk masker mereka dan dapat digunakan untuk meningkatkan kualitas produk serta layanan.

Kata Kunci: Analisis Sentimen; Naive Bayes; Data Mining; Ulasan Produk; E-Commerce

Abstract - This research analyzes sentiment in mask product reviews using the Naive Bayes algorithm. The dataset consists of 99 mask product reviews from an e-commerce platform, with positive and negative sentiment classifications. The methodology involves data preprocessing, modeling using the Naive Bayes algorithm, and model performance evaluation. The analysis results show that the Naive Bayes algorithm successfully classified review sentiments with an accuracy rate of 83.67%. This research provides valuable insights for e-commerce stakeholders in understanding customer perceptions of their mask products and can be used to improve product quality and service.

Keywords: Sentiment Analysis; Naive Bayes; Data Mining; Product Reviews; E-Commerce

1. PENDAHULUAN

Perkembangan e-commerce di Indonesia telah menghasilkan volume besar ulasan produk online, khususnya untuk produk kesehatan seperti masker yang menjadi kebutuhan utama masyarakat, terutama setelah pandemi COVID-19. Ulasan ini mengandung informasi berharga tentang persepsi dan kepuasan pelanggan yang dapat dimanfaatkan untuk meningkatkan kualitas produk dan layanan. Namun, dengan volume ulasan yang terus meningkat, menganalisis ulasan secara manual menjadi tidak efisien dan memakan waktu. Selain itu, ulasan pelanggan menggunakan bahasa yang beragam, baik dalam gaya maupun struktur kalimat, sehingga memerlukan metode analisis yang dapat menangani variasi tersebut secara efisien. Tantangan lain yang dihadapi adalah banyaknya ulasan yang mengandung noise atau ketidakteraturan, yang membuat analisis sentimen menjadi lebih kompleks dan membutuhkan solusi yang dapat mengatasi permasalahan tersebut.

Namun, menganalisis ulasan secara manual menjadi tidak efisien karena jumlahnya yang besar dan keragamannya dalam gaya bahasa. Sebagai alternatif, analisis sentimen berbasis teknik data mining menawarkan solusi otomatis yang lebih efisien. Salah satu algoritma yang sering digunakan untuk analisis sentimen adalah Naive Bayes. Algoritma ini terkenal karena kesederhanaannya dalam implementasi, kecepatan pemrosesan, dan kemampuan memberikan hasil akurat dalam klasifikasi teks meskipun dengan dataset yang terbatas. Menurut penelitian oleh (Marga, 2022), Naive Bayes menunjukkan performa yang konsisten dalam klasifikasi sentimen dengan akurasi hingga 85% pada dataset ulasan berbahasa Indonesia.

Penelitian ini bertujuan untuk memanfaatkan algoritma Naive Bayes untuk menganalisis sentimen ulasan produk masker di platform e-commerce. Dengan mengidentifikasi pola sentimen pelanggan, hasil analisis ini diharapkan dapat memberikan wawasan bagi penjual untuk meningkatkan layanan dan produk mereka, serta membantu konsumen dalam membuat keputusan pembelian yang lebih informatif. Selain itu, penelitian ini juga diharapkan dapat memberikan kontribusi teoretis dalam pengembangan teknologi analisis sentimen untuk teks berbahasa Indonesia.



Penelitian sebelumnya menunjukkan bahwa Naive Bayes efektif untuk klasifikasi teks pendek, seperti ulasan pelanggan, karena model probabilistiknya dapat mengatasi noise pada data (Yusuf, 2023). Dalam konteks ini, algoritma Naive Bayes diharapkan dapat mengatasi tantangan berupa variasi bahasa dan dataset yang terbatas.

2. METODE

2.1 Dataset dan Preprocessing

Dataset penelitian terdiri dari 99 ulasan produk masker dengan atribut: nama produk, akun pengguna, ulasan (yang telah dibersihkan), dan label sentimen (1 untuk positif, 0 untuk negatif). Preprocessing data meliputi:

- a. Case folding.
- b. Tokenization.
- c. Stopword removal.
- d. Stemming

2.2 Implementasi Naive Bayes

Implementasi menggunakan RapidMiner dengan tahapan:

- a. Pembagian data (70% training, 30% testing)
- b. Transformasi teks ke representasi vektor
- c. Pelatihan model Naive Bayes
- d. Validasi dan evaluasi performa

3. ANALISA DAN PEMBAHASAN

3.1 Karakteristik Dataset

Dari total 99 ulasan yang dianalisis, distribusi data menunjukkan:

- Sentimen Positif: 85 ulasan (85.86%)
- Sentimen Negatif: 14 ulasan (14.14%)

Karakteristik utama dataset:

- 1) Dominasi sentimen positif menunjukkan tingkat kepuasan pelanggan yang tinggi
- 2) Kata-kata yang sering muncul dalam ulasan positif: "bagus", "sesuai", "cepat", "mantap"
- 3) Kata-kata yang sering muncul dalam ulasan negatif: "kecewa", "lambat", "tipis", "putus"

3.2 Hasil Klasifikasi

Setelah implementasi algoritma Naive Bayes, diperoleh hasil:

- 1) Akurasi Model: 83.67%
- 2) Confusion Matrix:
 - True Positive (TP): 25
 - True Negative (TN): 4
 - False Positive (FP): 2
 - False Negative (FN): 3
- 3) Metrik Performa:

- Precision: 92.59%
- Recall: 89.29%
- F1-Score: 90.91%

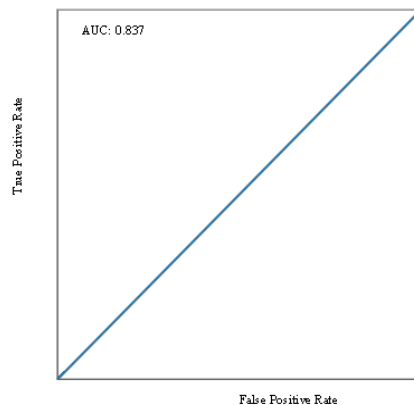
PerformanceVector

accuracy: 83.67%			
Class	precision	recall	f1_score
positive	92.59%	89.29%	90.91%
negative	66.67%	57.14%	61.54%

Confusion Matrix

	predicted positive	predicted negative
true positive	25	3
true negative	2	4

ROC Curve



Gambar 1. Visualisasi Data

3.3 Analisis Performa Model

1) Kelebihan Model:

- Mampu mengklasifikasikan sentimen positif dengan akurasi tinggi
- Precision tinggi menunjukkan ketepatan dalam prediksi sentimen positif
- Waktu komputasi yang efisien

2) Tantangan dan Keterbatasan:

- Dataset tidak seimbang (imbalanced) dengan dominasi sentimen positif
- Beberapa ulasan pendek sulit diklasifikasikan dengan tepat
- Kesalahan klasifikasi terutama terjadi pada ulasan dengan konteks ambigu

3.4 Implikasi Praktis

1) Untuk Penjual:

- a. Mayoritas ulasan positif menunjukkan kepuasan pelanggan yang tinggi



- b. Area perbaikan teridentifikasi dari analisis ulasan negatif:
 - Kualitas tali masker
 - Ketebalan masker
 - Kecepatan respon penjual
- 2) Untuk Platform E-commerce:
 - a. Model dapat digunakan untuk filtering ulasan otomatis
 - b. Membantu identifikasi early warning untuk masalah produk

4. KESIMPULAN

4.1 Kesimpulan

1. Algoritma Naive Bayes berhasil mengklasifikasikan sentimen ulasan produk masker dengan akurasi 83.67%
2. Performa model menunjukkan precision 92.59% dan recall 89.29%
3. Mayoritas pelanggan (85.86%) memberikan ulasan positif terhadap produk masker
4. Area perbaikan utama teridentifikasi pada kualitas produk dan layanan pelanggan

4.2 Saran

- 1) Pengembangan Model:
 - Penambahan jumlah dataset untuk meningkatkan akurasi
 - Implementasi teknik balancing data
 - Eksperimen dengan algoritma lain
- 2) Pengembangan Penelitian:
 - Analisis aspek spesifik produk
 - Implementasi deteksi sarkasme
 - Analisis temporal ulasan

REFERENCES

- Liu, B. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press.
- Pratama, B. Y., & Saputra, R. (2020). Analisis sentimen e-commerce di Indonesia menggunakan metode Naive Bayes. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 7(2), 109-119.
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093-1113. <https://doi.org/10.1016/j.asej.2014.04.011>
- Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. MIT Press.
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1–135. <https://doi.org/10.1561/15000000011>
- Marga, N. S. (2022). SENTIMEN ANALISIS TENTANG KEBIJAKAN PEMERINTAH TERHADAP KASUS CORONA MENGGUNAKAN METODE NAIVE BAYES. *Jurnal Informatika dan Rekayasa Perangkat Lunak*, 2(4), 453–463. <https://doi.org/10.33365/jatika.v2i4.1602>
- Yusuf, B. P. (2023). Analisis Sentimen Berbasis Aspek Terhadap Ulasan Aplikasi Market Place Menggunakan Bidirectional Long Short-Term Memory [Other, Universitas Komputer Indonesia]. https://doi.org/10.15.UNIKOM_BAGUS%20PERDANA%20YUSUF_BAB%205.pdf