



## Analisis Feedback Pengguna Aplikasi Shopee Menggunakan Distribusi Proporsi dan Teknologi Informasi

Dora Alvionita<sup>1</sup>, Lidiya<sup>2</sup>, Jadianan Parhusip<sup>3</sup>

<sup>123</sup>Universitas Palangka Raya, Palangka Raya, Kalimantan Tengah, Indonesia

Email: [doraalvionita@mhs.eng.upr.ac.id](mailto:doraalvionita@mhs.eng.upr.ac.id)<sup>1</sup>, [lidiya1997@mhs.eng.upr.ac.id](mailto:lidiya1997@mhs.eng.upr.ac.id)<sup>2</sup>,  
[parhusip.jadianan@it.upr.ac.id](mailto:parhusip.jadianan@it.upr.ac.id)<sup>3</sup>

Alamat Kampus: Jl. Hendrik Timang, Palangka Raya, Kalimantan Tengah 73111, Indonesia

Korespondensi penulis: [doraalvionita@mhs.eng.upr.ac.id](mailto:doraalvionita@mhs.eng.upr.ac.id)

**Abstract.** *This research aims to analyze Shopee application user feedback using Google Colab-based information technology and a proportion distribution approach. The method used includes a Multinomial Naïve Bayes-based classification technique to categorize sentiment into positive and negative. The research results show that the model accuracy level is 90.84% with an average precision for the positive and negative categories of 0.93 and 0.78 respectively. The statistical test produces a T-statistic value of 2759.50 with a p-value of 0.0, which shows significance in the difference in the distribution of positive and negative reviews.*

**Keywords:** *Classification, Proportion Distribution, Shopee, Information Technology, User Reviews.*

**Abstrak.** Penelitian ini bertujuan untuk menganalisis feedback pengguna aplikasi Shopee menggunakan teknologi informasi berbasis Google Colab dan pendekatan distribusi proporsi. Metode yang digunakan meliputi teknik klasifikasi berbasis Multinomial Naïve Bayes untuk mengkategorikan sentimen menjadi positif dan negatif. Hasil penelitian menunjukkan tingkat akurasi model sebesar 90,84% dengan rata-rata precision untuk kategori positif dan negatif masing-masing sebesar 0,93 dan 0,78. Uji statistik menghasilkan nilai T-statistik sebesar 2759,50 dengan p-value 0,0, yang menunjukkan signifikansi pada perbedaan distribusi ulasan positif dan negatif.

**Kata kunci:** Klasifikasi, Distribusi proporsi, Shopee, Teknologi Informasi, Ulasan Pengguna

### 1. LATAR BELAKANG

Aplikasi Shopee telah menjadi salah satu platform e-commerce terpopuler di Asia Tenggara. Seiring bertambahnya jumlah pengguna, ulasan dan feedback dari pengguna menjadi salah satu aspek penting yang dapat dimanfaatkan untuk meningkatkan kualitas layanan aplikasi. Namun, untuk mempertahankan kepuasan pelanggan, penting untuk memahami kebutuhan dan meningkatkan kualitas layanan. Penelitian ini bertujuan untuk menganalisis ulasan pengguna Shopee dengan pendekatan distribusi proporsi dan teknologi informasi berbasis Google Colab. Fokus utama adalah penerapan metode klasifikasi untuk sentimen ulasan dan evaluasi statistik menggunakan distribusi proporsi. Distribusi proporsi dipilih karena mampu menampilkan perbandingan secara visual terhadap dominasi sentimen positif dan negatif.

### 2. KAJIAN TEORITIS

#### Analisis Sentimen dan Metode Naïve Bayes

Analisis sentimen merupakan proses untuk memahami, mengekstraksi, dan mengolah data teks guna mendapatkan informasi sentiment yang terkandung di dalamnya, seperti opini positif atau negatif. Metode Naïve Bayes sering digunakan dalam analisis ini karena kesederhanaannya dan kemampuannya dalam menangani data teks. Penelitian oleh Wardani dan Sari menerapkan Naïve Bayes Classifier untuk analisis sentimen pada ulasan produk perawatan kulit wajah, dengan seleksi fitur N-gram dan Document *Frequency Thresholding*, yang meningkatkan akurasi klasifikasi [1]. Selain itu, penelitian oleh Wongkar dan Angdresey menggunakan algoritma Naïve Bayes untuk analisis sentimen data Twitter, menunjukkan efektivitas metode ini dalam mengklasifikasikan opini publik [2].

### **Distribusi Proporsi**

Distribusi proporsi adalah pendekatan statistik yang digunakan untuk menggambarkan persentase atau bagian dari suatu kelompok dalam populasi secara keseluruhan. Dalam analisis data, distribusi ini sering diterapkan untuk memahami pola distribusi dalam berbagai kategori, seperti dalam analisis sentimen, di mana proporsi ulasan positif dan negatif menjadi indikator penting kepuasan pelanggan. Studi oleh Universitas Telkom (2023) menjelaskan bahwa distribusi proporsi dapat memberikan wawasan mendalam tentang pola persepsi populasi, terutama dalam konteks evaluasi layanan berbasis data pelanggan. Pendekatan ini relevan untuk menampilkan pola distribusi sentimen secara visual dan kuantitatif, seperti yang diterapkan dalam penelitian ini untuk memahami dominasi sentimen positif dan negatif dalam dataset ulasan Shopee. [3].

### **Google Colab**

Google Collaboratory (Colab) adalah platform berbasis cloud yang mendukung penulisan dan eksekusi kode Python langsung di browser, tanpa memerlukan konfigurasi khusus. Colab menyediakan akses gratis ke sumber daya komputasi, termasuk GPU, yang memudahkan proses analisis data dan pengembangan model machine learning. Dokumentasi resmi Google Colab menjelaskan fitur-fitur dan cara penggunaannya [4]. Selain itu, panduan pembelajaran dari Universitas Bakrie menunjukkan penerapan R pada Google Colab untuk analisis statistik deskriptif, yang dapat diterapkan dalam berbagai disiplin ilmu [5].

## **3. METODE PENELITIAN**

Penelitian ini menggunakan dataset ulasan pengguna Shopee yang diambil dari Kaggle untuk menganalisis feedback atau umpan balik dari pengguna terhadap aplikasi Shopee. Dataset ini berisi berbagai ulasan yang diberikan oleh pengguna Shopee terkait pengalaman mereka dalam menggunakan aplikasi tersebut, yang mencakup aspek-aspek seperti kepuasan pengguna, kemudahan penggunaan, kualitas produk, serta aspek layanan dan pengiriman. Proses analisis dilakukan melalui beberapa tahap, dimulai dari pra-pemrosesan data, klasifikasi ulasan, evaluasi model, hingga analisis statistik. Pada tahap pra-pemrosesan data, data yang memiliki nilai kosong pada kolom teks ulasan dan rating ulasan dihapus untuk memastikan konsistensi analisis. Pembersihan teks ulasan dilakukan dengan menghapus karakter khusus, spasi berlebih, dan elemen yang tidak relevan. Selanjutnya, ditambahkan kolom sentimen berdasarkan rating ulasan dengan kategori positif untuk rating  $\geq 4$  dan negatif untuk rating  $< 4$ .

Pada tahap klasifikasi ulasan, algoritma Multinomial Naïve Bayes (MNB) digunakan untuk mengelompokkan ulasan menjadi dua kategori, yaitu positif dan negatif. Algoritma ini dipilih karena cocok untuk data teks dengan distribusi kata yang tidak merata. Sebelum proses klasifikasi, dilakukan vektorisasi teks menggunakan CountVectorizer untuk mengubah teks menjadi representasi numerik berbasis frekuensi kata. Model dievaluasi menggunakan metrik seperti akurasi, precision, recall, dan F1-score. Data dibagi menjadi 80% data latih dan 20% data uji menggunakan metode train-test split untuk memastikan validitas hasil evaluasi.

Tahap terakhir adalah analisis statistik menggunakan Uji T-Student untuk menguji hipotesis perbedaan distribusi antara ulasan positif dan negatif. Statistik uji dihitung berdasarkan distribusi rating ulasan di kedua kategori untuk menentukan apakah perbedaan signifikan secara statistik. Alur analisis data meliputi memuat dataset menggunakan pustaka Pandas, pembersihan data, vektorisasi teks, pelatihan model, evaluasi model, dan analisis statistik.

#### 4. HASIL DAN PEMBAHASAN

##### Hasil Analisis Data

Hasil analisis data menunjukkan bahwa model klasifikasi memiliki tingkat akurasi tinggi dengan nilai precision dan recall yang baik, terutama untuk kategori positif. Dari tabel ini, dapat disimpulkan bahwa mayoritas ulasan yang diberikan oleh pengguna adalah positif, menunjukkan kepuasan terhadap aplikasi Shopee. Namun, nilai precision dan recall yang lebih rendah untuk kategori negatif menandakan adanya potensi perbaikan dalam mengenali ulasan negatif. Hal ini penting untuk diinvestigasi lebih lanjut, mengingat ulasan negatif dapat memberikan wawasan berharga untuk meningkatkan layanan aplikasi.

Tabel 1. Evaluasi Model Klasifikasi

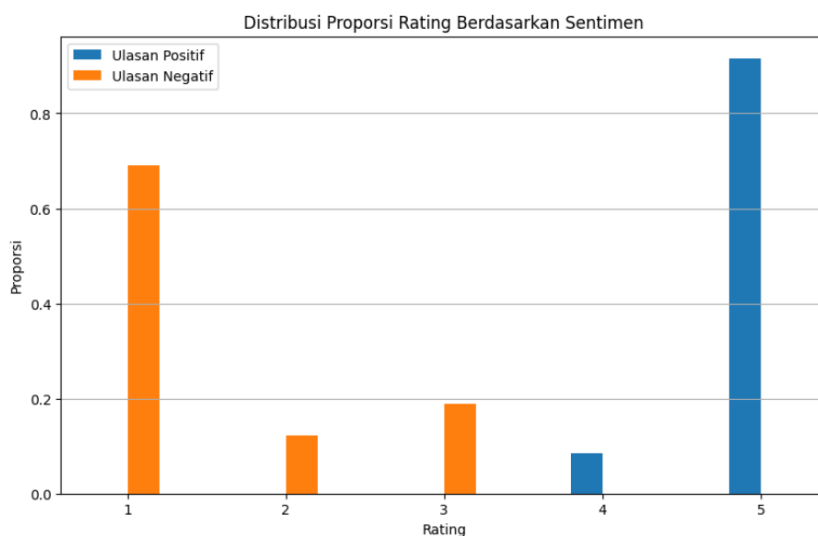
Kategori	Precision	Recall	F1-Score	Support
Negatif	0.78	0.67	0.72	27.399
Positif	0.93	0.96	0.95	128.969
<b>Total</b>	<b>0.90</b>	<b>0.91</b>	<b>0.91</b>	156.368

##### Uji Statistik

Hasil uji statistik menunjukkan bahwa nilai T-statistik adalah 2759,50 dengan p-value 0,0. Nilai statistik yang tinggi ini mencerminkan perbedaan yang sangat signifikan antara distribusi ulasan positif dan negatif. P-value sebesar 0,0 menunjukkan bahwa probabilitas untuk mendapatkan hasil seperti ini, jika tidak ada perbedaan nyata, hampir nol. Dengan demikian, hipotesis nol bahwa tidak ada perbedaan antara distribusi ulasan positif dan negatif dapat ditolak. Temuan ini mengindikasikan bahwa ulasan positif secara proporsi jauh lebih dominan dibandingkan ulasan negatif, mencerminkan kecenderungan umum pengguna untuk memberikan tanggapan yang lebih baik terhadap aplikasi Shopee. Hal ini menunjukkan bahwa terdapat perbedaan signifikan antara distribusi ulasan positif dan negatif.

##### Visualisasi Distribusi

Gambar berikut menunjukkan distribusi proporsi ulasan positif dan negatif berdasarkan rating:



Gambar 1. Grafik Distribusi Ulasan Positif dan Negatif

Dari grafik tersebut, terlihat bahwa:

- A. Ulasan positif mendominasi pada rating 4 dan 5.
- B. Ulasan negatif lebih sering muncul pada rating 1 hingga 3, menunjukkan bahwa ketidakpuasan pengguna umumnya dikaitkan dengan pengalaman buruk seperti pengiriman terlambat atau produk tidak sesuai.

## 5. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa distribusi proporsi dan teknologi informasi dapat digunakan untuk menganalisis feedback pengguna aplikasi Shopee. Hasilnya menunjukkan bahwa sebagian besar ulasan adalah positif, namun ulasan negatif tetap memberikan wawasan penting untuk perbaikan layanan. Sebagai contoh, ulasan negatif mencakup keluhan mengenai keterlambatan pengiriman atau ketidakcocokan produk yang diterima. Wawasan dari ulasan ini dapat diterapkan untuk meningkatkan proses logistik dan memastikan kesesuaian produk dengan deskripsi, yang pada akhirnya dapat meningkatkan kepuasan pelanggan secara keseluruhan. Penelitian lanjutan disarankan untuk mengeksplorasi korelasi antara fitur aplikasi dan ulasan pengguna.

## UCAPAN TERIMA KASIH

Ucapan terima kasih yang sebesar-besarnya kepada seluruh anggota tim yang telah berkontribusi dalam penyusunan jurnal ini. Setiap usaha, ide, dan kerja keras yang diberikan oleh rekan-rekan sangat berharga dalam menyelesaikan penelitian ini. Kerja sama yang baik dan dedikasi dari setiap individu menjadi kunci keberhasilan dalam proses ini. Semoga hasil yang dicapai dapat memberikan manfaat dan menjadi pijakan untuk penelitian lebih lanjut.

## DAFTAR REFERENSI

Wardani, S. K., & Sari, Y. A. (2021). Analisis Sentimen menggunakan Metode Naïve Bayes Classifier terhadap Review Produk Perawatan Kulit Wajah menggunakan

- Seleksi Fitur N-gram dan Document Frequency Thresholding. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 5(12), 5582–5590.
- Wongkar, M., & Angdresy, A. (2020). Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler: Twitter. *Jurnal Teknik Informatika*, 5(1), 1–8.
- Universitas Telkom. (2023). Distribusi Sampling Proporsi. Dalam *Panduan Pembelajaran Statistika*. Diakses dari [Telkom University Online Learning](#)
- Google. (2023). *Google Colab*. Diakses dari [Google Colab](#)
- Universitas Bakrie. (2020). Statistika Deskriptif menggunakan R pada Google Colab untuk Ilmu Komputer. Diakses dari [Bakrie Repository](#)