

ABSTRAKSI

Shopee merupakan salah satu aplikasi *e-commerce* dengan basis pengguna terbesar di Indonesia. Peningkatan jumlah penggunanya berbanding lurus dengan banyaknya ulasan yang ditinggalkan pada *Google Play Store*, sehingga ulasan tersebut dapat dijadikan sumber informasi penting mengenai pengalaman serta tingkat kepuasan pengguna. Penelitian ini dilakukan untuk menganalisis sentimen dari ulasan-ulasan tersebut menggunakan algoritma *Naïve Bayes*. Terdapat tiga fokus utama penelitian, yaitu: (1) menjelaskan tahapan *pre-processing* data teks ulasan, (2) menguji performa algoritma *Naïve Bayes* dalam mengklasifikasikan sentimen menjadi positif, negatif, dan netral, serta (3) menafsirkan hasil klasifikasi guna memberikan gambaran mengenai persepsi dan kepuasan pengguna terhadap aplikasi Shopee.

Data penelitian dikumpulkan melalui teknik web scraping pada ulasan aplikasi Shopee di *Google Play Store* dengan total 4.500 ulasan. Data dibagi dengan rasio 70% sebagai data latih dan 30% sebagai data uji. Tahapan *pre-processing* meliputi *cleaning*, *case folding*, *tokenizing*, *stopword removal*, normalisasi, serta *stemming*. Selanjutnya, data diberi label sesuai kategori sentimen untuk kemudian diproses menggunakan algoritma *Naïve Bayes*. Evaluasi kinerja model dilakukan dengan metrik akurasi, presisi, *recall*, dan *F1-score*.

Hasil pengujian menunjukkan bahwa algoritma *Naïve Bayes* mampu memberikan klasifikasi sentimen dengan tingkat akurasi yang baik. Mayoritas ulasan pengguna termasuk dalam kategori positif, yang mencerminkan citra dan kepuasan tinggi terhadap aplikasi Shopee. Penelitian ini diharapkan dapat menjadi masukan bagi pengembang aplikasi maupun pelaku bisnis *e-commerce* dalam merancang strategi peningkatan layanan berdasarkan persepsi pengguna.

Kata kunci: Analisis Sentimen, *Naïve Bayes*, Shopee, *Google Play Store*, Ulasan Pengguna.

ABSTRACT

Shopee is one of the largest e-commerce applications in Indonesia with a rapidly growing user base. Along with this growth, the number of reviews posted on Google Play Store also increases, making them a valuable source of information regarding user experiences and satisfaction levels. This study aims to analyze user review sentiments using the Naïve Bayes algorithm. The research focuses on three main aspects: (1) describing the text pre-processing stages of user reviews, (2) evaluating the performance of Naïve Bayes in classifying sentiments into positive, negative, and neutral categories, and (3) interpreting the classification results to provide insights into user perceptions and satisfaction toward the Shopee application.

The dataset consists of 4,500 reviews collected through web scraping on the Google Play Store. The data were split into 70% for training and 30% for testing. Pre-processing steps include cleaning, case folding, tokenizing, stopword removal, normalization, and stemming. The labeled data were then processed using the Naïve Bayes algorithm, and the model's performance was evaluated using accuracy, precision, recall, and F1-score metrics.

The results indicate that the Naïve Bayes algorithm achieved good accuracy in sentiment classification. Most user reviews fell into the positive category, reflecting favorable perceptions and high satisfaction toward the Shopee application. These findings are expected to provide valuable insights for application developers and e-commerce practitioners in designing strategies to improve services based on user opinions.

Keywords: *Sentiment Analysis, Naïve Bayes, Shopee, Google Play Store, User Reviews*

KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas Rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan tugas akhir berupa skripsi ini yang berjudul “ Analisis Sentimen pada Platform Shopee menggunakan Algoritma *Naïve Bayes*” dengan tepat waktu. Skripsi ini ditulis sebagai salah satu syarat untuk mendapatkan gelar Sarjana Strata-1 (S1) pada program studi Ilmu Komputer Universitas Putra Bangsa Kebumen.

Skripsi ini menjadi salah satu bentuk pencapaian penulis dalam mengaplikasikan ilmu yang diperoleh selama masa perkuliahan ke dalam praktik nyata, khususnya dalam kegiatan penelitian ilmiah. Proses penyusunan skripsi ini dilakukan dengan segala kemampuan yang dimiliki penulis, memadukan teori-teori yang relevan serta dukungan dari berbagai literatur yang menunjang. Penulis menyadari sepenuhnya bahwa karya ini masih jauh dari sempurna, mengingat keterbatasan ilmu dan pengalaman yang dimiliki.

Pada kesempatan yang berbahagia ini, penulis ingin menyampaikan rasa terima kasih yang sedalam-dalamnya kepada semua pihak yang telah memberikan dukungan, baik dalam bentuk bantuan nyata, dorongan semangat, maupun motivasi yang tidak ternilai harganya. Oleh karena itu, dengan penuh rasa hormat dan penghargaan, penulis menyampaikan terima kasih kepada:

1. Ibu Lolanda Hamim Annisa, S.Kom., M.Kom., selaku dosen pembimbing yang dengan penuh kesabaran, keikhlasan, serta ketelatenan telah membimbing, memberikan arahan, dan selalu memotivasi penulis hingga skripsi ini dapat terselesaikan.

2. Bapak Rahmat Hidayat, S.Kom., M.Kom., selaku dosen pembimbing di awal penyusunan skripsi, yang telah memberikan banyak masukan, bimbingan, serta ilmu berharga sebelum kemudian beliau berpindah tugas.
3. Bapak Awaludin A'bid, selaku Kepala Program Studi Ilmu Komputer Universitas Putra Bangsa, atas segala dukungan, arahan, serta fasilitas yang diberikan kepada penulis selama masa perkuliahan.
4. Seluruh dosen dan karyawan Universitas Putra Bangsa, yang telah memberikan ilmu pengetahuan, pengalaman, serta bantuan dalam berbagai bentuk selama penulis menempuh pendidikan hingga terselesaikannya skripsi ini.

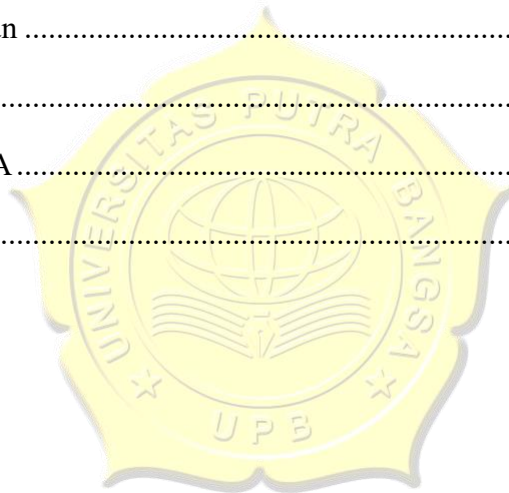
Penulis menyadari sepenuhnya bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan demi perbaikan di masa mendatang. Besar harapan penulis, semoga karya sederhana ini dapat memberikan manfaat, baik bagi penulis sendiri maupun bagi pembaca, serta dapat menjadi bagian kecil dari pengembangan ilmu pengetahuan.

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN UJIAN	iv
PERNYATAAN BEBAS PLAGIARISME	v
HALAMAN MOTTO	vi
HALAMAN PERSEMBAHAN.....	vii
ABSTRAKSI.....	ix
<i>ABSTRACT</i>	x
KATA PENGANTAR.....	xi
DAFTAR ISI	xiii
DAFTAR TABEL.....	xvi
DAFTAR GAMBAR	xvii
DAFTAR LAMPIRAN	xviii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Batasan Masalah.....	6
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian	7
BAB II KAJIAN PUSTAKA	9
2.1 Landasan Teori	9
2.1.1 Marketplace.....	9

2.1.2	Platform Shopee	10
2.1.3	<i>Machine Learning</i>	13
2.1.4	<i>Analisis Sentimen</i>	16
2.1.5	Algoritma <i>Naïve Bayes</i>	17
2.1.6	Python	19
2.1.7	<i>Web Scraping</i>	21
2.1.8	<i>Text Pre-processing</i>	22
2.2	Penelitian Terdahulu.....	24
BAB III METODE PENELITIAN.....		27
3.1	Jenis Penelitian.....	27
3.2	Tahap Penelitian	27
3.3	Sumber Data.....	36
3.4	Teknik Pengumpulan Data	41
3.4.1	Tahap Pengumpulan Data	41
3.5	Perangkat dan Alat yang Digunakan	43
3.6	Ilustrasi Hasil Akhir	46
BAB IV HASIL DAN PEMBAHASAN.....		49
4.1	Sample.....	49
4.1.1	<i>Scraping Data</i>	49
4.1.2	Eliminasi Data.....	51
4.2	<i>Pre-Processing Text</i>	53
4.2.1	<i>Cleaning</i>	53
4.2.2	<i>Case Folding</i>	54
4.2.3	<i>Tokenizing</i>	55
4.2.4	<i>Stopword Removal</i>	56

4.2.5	Normalisasi	57
4.2.6	<i>Stemming</i>	58
4.2.7	Pelabelan	59
4.2.8	<i>WordCloud</i>	60
4.3	Evaluasi Model.....	62
4.3.1	Pembagian Data Latih dan Data Uji.....	62
4.3.2	Implementasi Algoritma <i>Naïve Bayes</i>	62
4.3.3	Analisis Hasil Pengolahan Data	64
BAB V KESIMPULAN DAN SARAN.....		67
5.1	Kesimpulan	67
5.2	Saran.....	68
DAFTAR PUSTAKA		69
LAMPIRAN		73



DAFTAR TABEL

Tabel II- 1 Penelitian Terdahulu.....	24
---------------------------------------	----



DAFTAR GAMBAR

Gambar III-1 Alur Penelitian	28
Gambar III -2 Alur Pengumpulan Data.....	38
Gambar III- 3 Contoh Ulasan Rating	47
Gambar III- 4 Contoh WordCloud	48
Gambar IV- 1 Proses Scraping Data	50
Gambar IV- 2 Hasil Scraping Data	51
Gambar IV- 3 Proses Eliminasi Data	52
Gambar IV- 4 Hasil Eliminasi.....	52
Gambar IV- 5 Proses Cleaning.....	53
Gambar IV- 6 Hasil Cleaning.....	54
Gambar IV- 7 Proses Case Folding.....	54
Gambar IV- 8 Hasil Case Folding.....	55
Gambar IV- 9 Proses Tokenizing	55
Gambar IV- 10 Hasil Tokenizing	56
Gambar IV- 11 Proses Stopword Removal	56
Gambar IV- 12 Hasil Stopword Remover.....	57
Gambar IV- 13 Proses Normalisasi.....	57
Gambar IV- 14 Hasil Normalisasi.....	58
Gambar IV- 15 Proses Stemming.....	58
Gambar IV- 16 Hasil Stemming.....	59
Gambar IV- 17 Proses Labeling Data	60
Gambar IV- 18 Hasil Labeling Data	60
Gambar IV- 19 WordCloud Sentimen Positif.....	61
Gambar IV- 20 WordCloud Sentimen Negatif.....	61
Gambar IV- 21 WordCloud Sentimen Netral	61
Gambar IV- 22 Split Data.....	62
Gambar IV- 23 Hasil TF-IDF	63

DAFTAR LAMPIRAN

Lampiran I Sample Hasil Pre-Processing	74
Lampiran II Kartu Konsultasi Skripsi	75
Lampiran III Kartu Peserta Seminar	76

