

**ANALISIS SENTIMEN PENILAIAN PENGGUNA
MARKETPLACE LAZADA DENGAN METODE
NAÏVE BAYES DAN SUPPORT VECTOR MACHINE**

SKRIPSI



Diajukan Oleh:

SUGENG SETYABUDI

NPM. 21032010002

**PROGRAM STUDI TEKNIK INDUSTRI
FAKULTAS TEKNIK DAN SAINS
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR**

2025

**ANALISIS SENTIMEN PENILAIAN PENGGUNA
MARKETPLACE LAZADA DENGAN METODE
NAÏVE BAYES DAN SUPPORT VECTOR MACHINE**

SKRIPSI

Diajukan sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Program Studi Teknik Industri



Diajukan Oleh:
SUGENG SETYABUDI
NPM. 21032010002

**PROGRAM STUDI TEKNIK INDUSTRI
FAKULTAS TEKNIK DAN SAINS
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"
JAWA TIMUR
SURABAYA
2025**

SKRIPSI

**ANALISIS SENTIMEN PENILAIAN PENGGUNA
MARKETPLACE LAZADA DENGAN METODE
NAÏVE BAYES DAN SUPPORT VECTOR MACHINE**

Disusun Oleh:

SUGENG SETYABUDI

21032010002

Telah dipertahankan dihadapan Tim Penguji Skripsi dan diterima oleh
Publikasi Jurnal Akreditasi Sinta 1-3
Program Studi Teknik Industri Fakultas Teknik dan Sains
Universitas Pembangunan Nasional "Veteran" Jawa Timur Surabaya
Pada Tanggal : 7 Maret 2025

Tim Penguji :

1.

Ir. Sumiati, M.T.

NIP. 196012131991032001

Pembimbing :

1.

Enny Arvanny, S.T., M.T.

NIP. 197009282021212002

2.

Ir. Moch. Tutuk Saffrin, M.T.

NIP. 196304061989031001

Mengetahui,

**Dekan Fakultas Teknik dan Sains
Universitas Pembangunan Nasional "Veteran" Jawa Timur
Surabaya**

Prof. Dr. Dra. Jarivah, M.P.

NIP. 19650403 199103 2 001



KETERANGAN REVISI

Mahasiswa di bawah ini:

Nama : Sugeng Setyabudi
NPM : 21032010002
Program Studi : ~~Teknik Kimia / Teknik Industri / Teknologi Pangan /
Teknik Lingkungan / Teknik Sipil~~

Telah telah mengerjakan revisi / ~~tidak ada revisi *) PRA RENCANA (DESAIN) /
SKRIPSI / TUGAS AKHIR~~ Ujian Lisan Periode Maret, TA 2024/2025.

Dengan judul : **ANALISIS SENTIMEN PENILAIAN PENGGUNA
MARKETPLACE LAZADA DENGAN METODE NAÏVE
BAYES DAN SUPPORT VECTOR MACHINE**

Dosen yang memerintahkan revisi

1. Enny Aryanny, S.T., M.T.
2. Ir. Sumiati, M.T.
3. Ir. Moch. Tutuk Safirin, M.T.

Surabaya, 07 Maret 2025
Menyetujui,
Dosen Pembimbing

Enny Aryanny, S.T., M.T.
NIP. 197009282021212002

Catatan: *) coret yang tidak perlu



SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertanda tangan di bawah ini:

Nama : Sugeng Setyabudi
NPM : 21032010002
Program : Sarjana (S1)
Program Studi : Teknik Industri
Fakultas : Teknik dan Sains

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.

Surabaya, 11 Maret 2025

Yang Membuat Pernyataan



Sugeng Setyabudi
NPM. 21032010002

KATA PENGANTAR

Puji syukur kepada Allah SWT berkat Rahmat, Hidayah, dan Karunia-Nya kepada kita semua sehingga penulis dapat menyelesaikan skripsi dengan judul **“ANALISIS SENTIMEN PENILAIAN PENGGUNA *MARKETPLACE LAZADA* DENGAN METODE *NAÏVE BAYES* DAN *SUPPORT VECTOR MACHINE*”**. Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan program Sarjana Strata 1 (S1) Program Studi Teknik Industri Fakultas Teknik dan Sains, Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Penulis menyadari bahwa penyusunan skripsi ini tidak akan terselesaikan tanpa dukungan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Ir. Akhmad Fauzi, MMT., IPU, selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Ibu Prof. Dr. Dra. Jariyah, M.P., selaku Dekan Fakultas Teknik dan Sains Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Bapak Ir. Rusindiyanto, M.T., selaku Koordinator Program Studi Teknik Industri UPN Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Ibu Enny Aryanny, S.T., M.T. selaku dosen pembimbing, atas bimbingan, saran, dan motivasi yang telah diberikan.
5. Ibu dan Bapak Dosen Penguji yang telah memberikan koreksi, saran, dan arahan dalam penyempurnaan Skripsi ini.

6. Seluruh dosen Teknik Industri yang telah berbagi ilmu selama masa perkuliahan.
7. Orang tua dan keluarga penulis yang selalu mengantarkan doa, dukungan, serta kasih sayang yang selalu tercurah selama ini.
8. Seluruh dosen Teknik Industri yang telah berbagi ilmu selama masa perkuliahan.
9. Teman-teman Teknik Industri angkatan 2021 atas dukungan dan motivasi yang terus diberikan.
10. Tak lupa, Eko dan Ocil, hewan peliharaan saya yang tanpa mereka sadari, telah menemani saya melewati banyak momen selama penyusunan skripsi ini.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan. Oleh karena itu, penulis sangat mengharapkan saran dan kritik yang membangun untuk perbaikan lebih lanjut, dengan harapan skripsi ini dapat bermanfaat dalam bidang pendidikan dan penerapannya di lapangan serta dapat dikembangkan lebih jauh.

Surabaya, 03 Maret 2025

Penulis

DAFTAR ISI

KATA PENGANTAR	i
DAFTAR ISI	iii
DAFTAR TABEL	v
DAFTAR GAMBAR	vi
ABSTRAK	vii
ABSTRAC	viii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	5
1.4 Asumsi-Asumsi.....	5
1.5 Tujuan	5
1.6 Manfaat Penelitian.....	6
1.7 Sistematika Penulisan.....	7
BAB II TINJAUAN PUSTAKA	9
2.1 Analisis Sentimen	9
2.2 <i>Marketplace</i>	10
2.3 <i>Web Scraping</i>	10
2.3.1 Google Play Scraper.....	11
2.4 <i>Text Mining</i>	12
2.5 Pelabelan Data	13
2.6 <i>Text Preprocessing</i>	14
2.7 Pembobotan Kata <i>Term Frequency - Inverse Doc Frequency</i>	17
2.8 Pembagian Data Latih dan Data Uji	18
2.9 <i>Data Mining</i>	19
2.9.1 Klasifikasi.....	20
2.10 <i>Naïve Bayes</i>	21
2.11 <i>Support Vector Machine</i>	23
2.12 Evaluasi Model Klasifikasi.....	25

2.13	<i>Word Cloud</i>	28
2.14	5W+1H	29
2.15	Penelitian Terdahulu	31
BAB III METODOLOGI PENELITIAN		38
3.1	Lokasi dan Tempat Penelitian.....	38
3.2	Identifikasi dan Definisi Operasional variabel	38
3.2.1	Variabel Terikat	38
3.2.2	Variabel Bebas.....	39
3.3	Langkah-langkah Pemecahan Masalah	39
BAB IV HASIL DAN PEMBAHASAN		46
4.1	Pengumpulan Data	46
4.2	Pengolahan Data	47
4.2.1	Pelabelan Data	48
4.2.2	Tahap <i>Text Preprocessing</i>	50
4.2.3	Pembobotan Kata <i>Term Frequency - Inverse Document Frequency</i>	55
4.2.4	Pembagian Data.....	59
4.2.5	Klasifikasi <i>Naïve Bayes</i>	60
4.2.6	Klasifikasi <i>Support Vector Machine</i>	62
4.2.7	Perbandingan Metode.....	64
4.2.8	Visualisasai <i>Word Cloud</i>	65
4.2.9	Usulan Perbaikan dengan Metode 5W+1H.....	72
4.3	Hasil Dan Pembahasan	84
BAB V KESIMPULAN DAN SARAN		86
5.1	Kesimpulan.....	86
5.2	Saran.....	87
DAFTAR PUSTAKA.....		88
LAMPIRAN.....		98

DAFTAR TABEL

Tabel 2.1 Contoh Pelabelan Data Berdasarkan Rating	14
Tabel 2.2 Contoh Case Folding.....	15
Tabel 2.3 Contoh <i>Tokenizing</i>	16
Tabel 2.4 Contoh <i>Filtering Stopword</i>	16
Tabel 2.5 Contoh <i>Stemming</i>	17
Tabel 2.6 <i>Confussion Matrix</i>	26
Tabel 2.7 Contoh Analisis 5W+1H	30
Tabel 4.1 Hasil Pengumpulan Data.....	46
Tabel 4.2 Hasil Dari Pemberian Label	48
Tabel 4.3 Hasil Dari <i>Cleaning Data</i>	51
Tabel 4.4 Hasil Dari <i>Case Folding</i>	52
Tabel 4.5 Hasil Dari Stopword Removal	52
Tabel 4.6 Hasil Dari <i>Tokenizing</i>	54
Tabel 4.7 Hasil Dari <i>Stemming</i>	55
Tabel 4.8 Contoh Salah Satu Komentar Yang Memiliki Kata ‘barang’	57
Tabel 4.9 Bobot Kata Yang Sering Muncul Dalam Semua Komentar	58
Tabel 4.10 <i>Confussion Matrix Naïve Bayes</i>	62
Tabel 4.11 <i>Confussion Support Vector Machine</i>	64
Tabel 4.12 Hasil Akurasi	65
Tabel 4.13 Contoh Komentar Sentimen Negatif Yang Menyoroti Kekurangan Aplikasi Lazada.....	72
Tabel 4.14 Usulan Perbaikan Dengan Metode 5W+1H Pada Kekurangan Aplikasi Lazada	73
Tabel 4.15 Contoh Komentar Sentimen Negatif Tentang Masalah Pengiriman ...	74
Tabel 4.16 Usulan Perbaikan Dengan Metode 5W+1H Terhadap Masalah Pengiriman	75
Tabel 4.17 Contoh Komentar Sentimen Negatif Terhadap Masalah Barang	77
Tabel 4.18 Usulan Perbaikan Dengan Metode 5W+1H Terhadap Masalah Barang	78
Tabel 4.19 Contoh Komentar Sentimen Negatif Terhadap Sistem Pembayaran Buruk.....	79
Tabel 4.20 Usulan Perbaikan Dengan Metode 5W+1H Terhadap Masalah Pembayaran.....	80
Tabel 4.21 Contoh Komentar Sentimen Negatif Terhadap Pengalaman Belanja Yang Buruk	82
Tabel 4.22 Usulan Perbaikan Dengan Metode 5W+1H Terhadap Pengalaman Belanja Yang Buruk	83

DAFTAR GAMBAR

Gambar 1.1 Kunjungan 5 <i>Marketplace</i> 2023.....	2
Gambar 1.2 Perolehan Rating Bintang Aplikasi Lazada Selama Oktober 2024.....	3
Gambar 2.1 Contoh Visualisasi <i>Word Cloud</i>	29
Gambar 3.1 Langkah-langkah Pemecahan Masalah.....	40
Gambar 4.1 Data Jumlah Sentimen Positif, Netral Dan Negatif.....	49
Gambar 4.2 Lima Kata Yang Paling Sering Muncul Pada Komentar Lazada	56
Gambar 4.3 Hasil Pembagian Data Latih Dan Data Uji	59
Gambar 4.4 Hasil Evaluasi Kinerja <i>Naïve Bayes</i>	61
Gambar 4.5 Hasil Klasifikasi <i>Support Vector Machine</i>	63
Gambar 4.6 <i>Word Cloud</i> Sentimen Positif.....	66
Gambar 4.7 Kata Yang Paling Sering Muncul Pada Sentimen Positif.....	67
Gambar 4.8 <i>Word Cloud</i> Sentimen Netral.....	68
Gambar 4. 9 Kata Yang Paling Sering Muncul Pada Sentimen Netral	69
Gambar 4.10 <i>Word Cloud</i> Sentimen Negatif.....	70
Gambar 4.11 Kata Yang Paling Sering Muncul Pada Sentimen Negatif	71

ABSTRAK

Metode *Naïve Bayes* dan *Support Vector Machine* (SVM) digunakan dalam penelitian ini untuk menganalisis persepsi pengguna terhadap pasar Lazada. Untuk meningkatkan layanan dan strategi bisnis, penting untuk memahami perasaan pelanggan. Data dikumpulkan melalui *scraping web*, yang menghasilkan 5.261 komentar yang dibagi menjadi dua puluh persen data uji dan delapan puluh persen data latih. Tahapan preprocessing mencakup pembersihan teks, normalisasi huruf, penghapusan kata tidak penting, tokenisasi, dan stemming agar data lebih terstruktur. Hasil analisis menunjukkan bahwa SVM memiliki akurasi lebih tinggi, yakni 75%, dibandingkan dengan *Naïve Bayes* yang mencapai 72%. *Confusion matrix* mengonfirmasi bahwa SVM lebih akurat dalam mengklasifikasikan sentimen positif, netral, dan negatif. Dari visualisasi *word cloud*, ditemukan bahwa keluhan utama pengguna berkaitan dengan aplikasi yang sering bermasalah, keterlambatan pengiriman, produk yang tidak sesuai dengan deskripsi, kendala dalam proses pembayaran, serta pengalaman belanja yang kurang memuaskan. Hasil penelitian ini memberikan wawasan bagi Lazada untuk meningkatkan kualitas layanan dengan memperbaiki stabilitas aplikasi, mengoptimalkan logistik, memastikan keakuratan produk yang dijual, serta menyederhanakan proses pembayaran. Dengan perbaikan ini, diharapkan tingkat kepuasan dan loyalitas pelanggan terhadap Lazada dapat meningkat.

Kata Kunci - Analisis Sentimen, *Marketplace*, *Naïve Bayes*, *Support Vector Machine*.

ABSTRAC

Naïve Bayes and Support Vector Machine (SVM) methods are used in this study to analyze user perception of Lazada marketplace. To improve services and business strategies, it is important to understand customer feelings. Data was collected through web scraping, which resulted in 5,261 comments divided into twenty percent test data and eighty percent training data. Preprocessing stages included text cleaning, font normalization, removal of unimportant words, tokenization, and stemming to make the data more structured. The analysis results show that SVM has a higher accuracy of 75%, compared to Naïve Bayes which reaches 72%. Confusion matrix confirms that SVM is more accurate in classifying positive, neutral, and negative sentiments. From the word cloud visualization, it was found that the main complaints of users were related to frequent application issues, delivery delays, products that did not match the description, problems in the payment process, and an unsatisfactory shopping experience. The results of this study provide insights for Lazada to improve service quality by improving application stability, optimizing logistics, ensuring the accuracy of products sold, and simplifying the payment process. With these improvements, it is expected that the level of customer satisfaction and loyalty to Lazada can increase.

Keywords - Marketplace, Naïve Bayes, Sentiment Analysis, Support Vector Machine