

**ANALISIS SENTIMEN ULASAN PRODUK IPHONE
SECOND PADA TOKOPEDIA MENGGUNAKAN
METODE *NAÏVE BAYES* DAN *SUPPORT
VECTOR MACHINE***

SKRIPSI

**Oleh
Krisna Novianto
202110715003**



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2025**

LEMBAR PERSETUJUAN PEMBIMBING
(DIGUNAKAN UNTUK TUGAS AKHIR)

Judul Tugas Akhir : Analisis Sentimen Ulasan Produk iPhone *Second*
Pada Tokopedia Menggunakan Metode *Naïve*
Bayes dan *Support Vector Machine*

Nama Mahasiswa : Krisna Novianto

Nomor Pokok Mahasiswa : 202110715003

Program Studi/Fakultas : Informatika / Ilmu Komputer

Bekasi, 12 Juni 2025

MENYETUJUI

Pembimbing I



Dr. Herlawati, S.Si., M.M., M.Kom.

NIDN : 0311097302

Pembimbing II



Agus Hidayat, S.Kom., M.Si.

NIDN : 0304088303

Ketua Program Studi



Ahmad Fathurrozi, S.E., M.M.S.I

NIP: 2012486

Program Studi Informatika
Fakultas Ilmu Komputer
Universitas Bhayangkara Jakarta Raya
2025

LEMBAR PENGESAHAN

Judul Tugas Akhir : Analisis Sentimen Ulasan Produk iPhone *Second*
Pada Tokopedia Menggunakan Metode *Naïve*
Bayes dan *Support Vector Machine* (SVM)

Nama Mahasiswa : Krisna Novianto

Nomor Pokok Mahasiswa : 202110715003

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Tugas Akhir : 7 Juli 2025

Bekasi, 18 Juli 2025

MENGETAHUI,

Ketua Tim Penguji : Adi Muhajirin, S.Kom., M.Kom., M.M.

NIDN : 0318038501

Penguji I : Dr. Rakhmi Khalida, S.T., M.M.S.I.

NIDN : 0304099201

Penguji II : Dr. Herlawati, S.Si., M.M., M.Kom.

NIDN : 0311097302

MENGETAHUI,

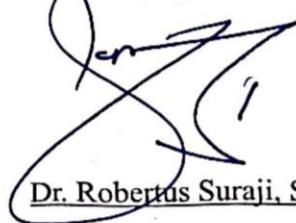
Ketua
Program Studi Informatika



Ahmad Fathurrozi, S.E., M.M.SI

NIP: 2012486

Dekan
Fakultas Ilmu Komputer



Dr. Robertus Suraji, S.S., M.A

NIP: 1904406



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Krisna Novianto
NPM : 202110715003
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Sentimen Ulasan Produk iPhone Second pada Tokopedia Menggunakan Metode *Naïve Bayes* dan *Support Vector Machine*

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggung jawabkannya sekaligus terhadap karya orang lain, maka saya bersedia menerima saksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya. Analisis Sentimen Ulasan Produk iPhone *Second* Pada Tokopedia Menggunakan Metode *Naïve Bayes* dan *Support Vector Machine*

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun

Bekasi, 5 Juni 2025



KRISNA NOVIANTO

ABSTRAK

Krisna Novianto, 202110715003. Analisis Sentimen Ulasan Produk iPhone *Second* Pada Tokopedia Menggunakan Metode *Naïve Bayes* dan *Support Vector Machine*. Bekasi: Fakultas Ilmu Komputer. Universitas Bhayangkara Jakarta Raya.

Tingginya harga iPhone baru mendorong sebagian masyarakat, khususnya dari kalangan menengah ke bawah, untuk beralih ke produk iPhone *second*. Tokopedia menjadi salah satu marketplace yang banyak dipilih karena reputasinya yang dianggap terpercaya. Berdasarkan survei terhadap 30 responden, sebanyak 91% menyatakan pernah membeli iPhone *second* di Tokopedia, dan 58,1% memilih platform ini karena alasan kepercayaan. Meski demikian, konsumen masih menghadapi sejumlah risiko, seperti ketidaksesuaian kondisi produk, kerusakan pada komponen penting (seperti baterai dan sinyal), hingga masalah legalitas seperti IMEI yang tidak tercatat resmi di bea cukai. Mayoritas informasi ini baru diketahui setelah membaca ulasan dari pengguna sebelumnya. Penelitian ini bertujuan untuk menganalisis sentimen dari 1.863 ulasan terhadap produk iPhone *second* di Tokopedia guna memberikan wawasan mengenai pengalaman konsumen. Dua algoritma klasifikasi *Naïve Bayes* dan *Support Vector Machine* (SVM)—digunakan untuk memproses data tersebut. Hasil evaluasi menunjukkan bahwa SVM memperoleh performa tertinggi dengan akurasi mencapai 96%, sedangkan *Naïve Bayes* hanya mencapai 93%. Analisis lanjutan menunjukkan bahwa ulasan positif umumnya menyoroti kualitas fisik dan kecepatan pengiriman, sementara ulasan negatif lebih sering menyinggung persoalan teknis dan keaslian perangkat. Model terbaik diintegrasikan ke dalam aplikasi berbasis *web* yang dilengkapi fitur prediksi manual dan visualisasi *word cloud*. Berdasarkan hasil uji pengguna, aplikasi ini dinilai mampu membantu calon pembeli dalam memahami kualitas produk secara lebih cepat dan efisien. Penelitian ini membuktikan bahwa pendekatan analisis sentimen berbasis *machine learning* dapat menjadi alat pendukung dalam pengambilan keputusan pembelian produk bekas secara daring.

Kata Kunci : iPhone *Second*, Tokopedia, Sentimen, *Naïve Bayes*, SVM

ABSTRACT

Krisna Novianto, 202110715003. *Sentiment Analysis of Second-Hand iPhone Reviews on Tokopedia Using Naïve Bayes and Support Vector Machine. Bekasi: Faculty of Computer Science. University of Bhayangkara Jakarta Raya*

The high price of new iPhones has driven a portion of society, particularly those from the middle to lower economic classes, to turn to second-hand iPhones. Tokopedia is one of the most frequently chosen marketplaces due to its reputation for trustworthiness. Based on a survey of 30 respondents, 91% stated that they had purchased a second-hand iPhone on Tokopedia, and 58.1% cited trust as the main reason for choosing the platform. However, consumers still face several risks, such as discrepancies in product condition, damage to essential components (such as batteries and signal reception), and legal issues like IMEI numbers not being officially registered with customs. Most of this information is only discovered after reading reviews from previous users. This study aims to analyze the sentiment of 1,863 product reviews of second-hand iPhones on Tokopedia to gain insights into consumer experiences. Two classification algorithms Naïve Bayes and Support Vector Machine (SVM)—were used to process the data. Evaluation results show that SVM achieved the highest performance with an accuracy of 96%, while Naïve Bayes reached only 93%. Further analysis reveals that positive reviews typically highlight physical condition and fast delivery, while negative reviews often mention technical issues and concerns about product authenticity. The best-performing model was integrated into a web-based application equipped with manual prediction and word cloud visualization features. Based on user testing, the application is considered helpful for prospective buyers in assessing product quality more quickly and efficiently. This research demonstrates that sentiment analysis using machine learning can be a supportive tool in online second-hand product purchasing decisions.

Keywords: *Second-hand iPhone, Tokopedia, Sentiment, Naïve Bayes, SVM*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : KRISNA NOVIANTO
NPM : 202110715003
Program Studi : INFORMATIKA
Fakultas : ILMU KOMPUTER
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (Non-Exclusive Royalty-Free Right)**, atas karya ilmiah saya yang berjudul :

ANALISIS SENTIMEN ULASAN PRODUK IPHONE SECOND PADA
TOKOPEDIA MENGGUNAKAN METODE NAIVE BAYES DAN SUPPORT
VECTOR MACHINE

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi
Pada tanggal : 18 Juli 2025
Yang Menyatakan



KRISNA NOVIANTO

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan Rahmat serta karunia-Nya, sehingga penulis bisa menyelesaikan proposal skripsi ini. Proposal skripsi ini berjudul *Analisis Sentimen Ulasan Produk iPhone Second Pada Tokopedia Menggunakan Metode Naïve Bayes dan Support Vector Machine* yang disusun sebagai prasyarat agar bisa mencapai gelar Sarjana Strata Satu (S1) pada Program Studi Informatika Universitas Bhayangkara Jakarta Raya. Penulis menyadari kekurangan dalam melakukan penulisan ini dan jauh dari kata sempurna.

Pada kesempatan ini penulis berterimakasih kepada seluruh pihak yang membantu serta moril maupun materil, serta jujur kepada:

1. Bapak Irjen Pol (Purn) Dr. Drs. Bambang Karsono, S. H., M. M., selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Bapak Dr. Robertus Suraji, S.S., M.A selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fathurrozi, S.E., MMSI. Selaku Ketua Program Studi Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya
4. Ibu Dr. Herlawati, S.Si., M.M., M.Kom., selaku Dosen Pembimbing I Proposal Skripsi yang memberikan masukan serta saran dan motivasi dalam mengerjakan tugas akhir ini.
5. Bapak Agus Hidayat, S. Kom., M. Si., selaku Dosen Pembimbing II Proposal Skripsi yang memberikan masukan serta saran dan motivasi dalam mengerjakan tugas akhir ini.

6. Ibu Dwipa Handayani, S. Kom., MMSI., selaku Dosen Pembimbing Akademik yang memberikan masukan serta motivasi dalam proses pengerjaan proposal skripsi.
7. Seluruh Dosen, staff dan mahasiswa Universitas Bhayangkara Jakarta Raya yang telah membantu dalam pengisian kuesioner guna mempertajam permasalahan pada penelitian saya terhadap produk *iPhone second*.
8. Kedua orang tua yang selalu mendoakan, memberi semangat serta mendukung dalam pekerjaan penelitian ini.
9. Teman-teman seperjuangan, sahabat, dan setiap orang entah sengaja atau tidak sengaja saya temui untuk berbagi cerita dan membantu setiap perjuangan saya dalam menjalani skripsi ini, terima kasih atas waktu yang diberikan.

Penulis menyadari bahwa penulisan proposal skripsi ini masih memiliki kekurangan, Oleh karena itu saran dan kritik sangat diharapkan untuk perbaikan di masa mendatang. Akhir, semoga proposal skripsi ini dapat bermanfaat bagi rekan-rekan mahasiswa/i serta para pembaca sekaligus demi menambah pengetahuan tentang proposal skripsi.

Bekasi, 17 Maret 2025

Penulis

Krisna Novianto

202110715003

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	8
1.3 Rumusan Masalah	9
1.4 Tujuan Penelitian	9
1.5 Manfaat Penelitian	10
1.6 Batasan Masalah.....	11
1.7 Sistematika Penulisan	12
BAB II TINJAUAN PUSTAKA.....	14
2.1 Penelitian Terdahulu	14
2.2 Teori Pendukung	30
2.2.1 <i>Machine Learning</i>	31
2.2.2 <i>Data Mining</i>	32
2.2.3 CRISP-DM (<i>Cross Industry Standard Process for Data Mining</i>)....	33
2.2.4 <i>Text Mining</i>	35
2.2.5 Analisis Sentimen	35
2.2.6 <i>Web Scraping</i>	36
2.2.7 <i>Naïve Bayes</i>	36
2.2.8 <i>Support Vector Machine (SVM)</i>	38
2.2.9 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	40
2.2.10 VADER.....	42

2.2.11	<i>Selenium</i>	42
2.2.12	<i>BeautifulSoup</i>	43
2.2.13	<i>Pandas</i>	43
2.2.14	<i>Confusion Metrix</i>	43
2.2.15	<i>Synthetic Minority Over-sampling Technique (SMOTE)</i>	46
2.2.16	<i>Custom Lexicon</i>	47
2.2.17	<i>Word Cloud</i>	47
2.2.18	<i>Klasifikasi</i>	47
2.2.19	<i>Stemming</i>	48
2.2.20	<i>Python</i>	48
2.2.21	<i>iPhone</i>	49
2.2.22	<i>Tokopedia</i>	50
2.2.23	<i>Visual Studio Code</i>	50
2.2.24	<i>Google Colab (Collaboratory)</i>	51
BAB III	METODOLOGI PENELITIAN	52
3.1	<i>Objek Penelitian</i>	52
3.2	<i>Kerangka Penelitian</i>	56
3.2.1	<i>Business Understanding</i>	57
3.2.2	<i>Data Understanding</i>	58
3.2.3	<i>Data Preparation</i>	60
3.2.4	<i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	64
3.2.5	<i>Modeling</i>	65
3.2.6	<i>Evaluation</i>	69
3.2.7	<i>Deployment</i>	71
3.3	<i>Kebutuhan Sistem Aplikasi Web</i>	78
3.4	<i>Metode Pengumpulan Data</i>	81
3.5	<i>Analisis Kebutuhan Sistem</i>	82
BAB IV	HASIL DAN PEMBAHASAN	84
4.1	<i>Pengumpulan Data</i>	84
4.2	<i>Business Understanding</i>	88
4.3	<i>Data Understanding</i>	89
4.4	<i>Data Preparation</i>	91
4.5	<i>Term Frequency (TF) dan Inverse Document Frequency (IDF)</i>	112
4.6	<i>Text Mining</i>	118

4.7	<i>Evaluation</i>	132
4.8	<i>Deployment</i>	142
BAB V PENUTUP		154
5.1	Kesimpulan	154
5.2	Saran.....	155
DAFTAR PUSTAKA		159
LAMPIRAN		167



DAFTAR TABEL

Tabel 2. 1 Tabel Analisis Hasil dan Evaluasi Studi Terkait	14
Tabel 4. 1 Perbandingan Kata Sebelum dan Setelah Proses Normalisasi.....	94
Tabel 4. 2 Daftar Kata-Kata dengan Sentimen Negatif	99
Tabel 4. 3 Daftar Kata Ulasan ke-1	100
Tabel 4. 4 Daftar Kata Ulasan ke-2.....	101
Tabel 4. 5 Kata-Kata Penting yang dipertahankan pada Proses <i>Stopword</i>	105
Tabel 4. 6 Perbandingan Ulasan Sebelum dan Setelah Proses <i>Stopword</i>	105
Tabel 4. 7 Perbandingan Ulasan Sebelum dan Setelah Proses <i>Stemming</i>	107
Tabel 4. 8 <i>Top</i> 10 Kata Sentimen Positif	110
Tabel 4. 9 <i>Top</i> 10 Kata Sentimen Negatif.....	112
Tabel 4. 10 Sebagian Ulasan Hasil <i>Preprocessing</i>	115
Tabel 4. 11 Hasil Perhitungan <i>TF-IDF</i>	116
Tabel 4. 12 Data <i>Training</i>	121
Tabel 4. 13 Hasil Perhitungan Probabilitas Pada Ulasan 1	125
Tabel 4. 14 Hasil Perhitungan Probabilitas Pada Ulasan 2	125
Tabel 4. 15 Ulasan 3 dan Ulasan 4 untuk Data Pengujian	126
Tabel 4. 16 Ulasan Untuk Data Pengujian	128
Tabel 4. 17 Data Training	130
Tabel 4. 18 Perbandingan evaluasi kedua model yaitu <i>Naïve Bayes</i> dan SVM..	140

DAFTAR GAMBAR

Gambar 1. 1 Tren Penurunan Total Kunjungan Tokopedia 2025	2
Gambar 1. 2 Ulasan Toko Boboy Store di Tokopedia	3
Gambar 1. 3 Preferensi <i>Marketplace</i> untuk Pembelian iPhone <i>second</i>	7
Gambar 1. 4 Distribusi Alasan Konsumen Membeli iPhone <i>second</i> di Tokopedia	7
Gambar 2. 1 Posisi <i>Machine Learning</i> dalam Ranah Kecerdasan Buatan.....	32
Gambar 2. 2 Tahapan pada <i>CRISP-DM</i>	33
Gambar 2. 3 <i>Confusion Matrix</i>	44
Gambar 2. 4 Tampilan iPhone 16 <i>Pro</i>	50
Gambar 2. 5 Logo Aplikasi Tokopedia.....	51
Gambar 2. 6 Logo Aplikasi <i>Visual Studio Code</i>	51
Gambar 2. 7 Logo <i>Google Studio Code</i>	52
Gambar 3. 1 Tampilan Halaman Toko Boboy Store di Tokopedia	54
Gambar 3. 2 Tampilan Halaman Toko JoyTaroGadgetOfficial di Tokopedia	54
Gambar 3. 3 Tampilan Halaman Toko Maxsense Store di Tokopedia	55
Gambar 3. 4 Tampilan Halaman Toko Yunishop00 di Tokopedia.....	56
Gambar 3. 5 Kerangka Penelitian	57
Gambar 3. 6 Hasil <i>Scraping</i> Ulasan pada Toko Boboystore	60
Gambar 3. 7 Produk Bukan iPhone <i>Second</i>	60
Gambar 3. 8 Hasil Penggabungan Dataset.....	61
Gambar 3. 9 Alur Tahapan Pemodelan Algoritma <i>Naïve Bayes (Multinomial)</i> ...	67
Gambar 3. 10 Alur Tahapan Pemodelan Algoritma <i>Support Vector Machine</i>	69

Gambar 3. 11 Activity Diagram Fitur Uji Model SVM.....	72
Gambar 3. 12 Activity Diagram Fitur Uji Model <i>Naïve Bayes</i>	74
Gambar 3. 13 Activity Diagram Fitur Visualisasi	76
Gambar 3. 14 <i>Activity Diagram</i> Fitur Tentang	78
Gambar 4. 1 Ulasan Konsumen Produk iPhone <i>Second</i>	84
Gambar 4. 2 Instalasi <i>Packages</i> atau <i>Libray</i>	85
Gambar 4. 3 Proses Penyaringan Ulasan Berdasarkan Produk iPhone Bekas	86
Gambar 4. 4 Hasil Penyaringan Produk.....	86
Gambar 4. 5 Proses Penggabungan Data	88
Gambar 4. 6 Informasi Dataset Ulasan iPhone <i>Second</i>	90
Gambar 4. 7 Perbandingan Kolom Ulasan Sebelum dan Sesudah Pembersihan..	92
Gambar 4. 8 Kode Penghapusan Kolom dan Perubahan Nama Kolom.....	93
Gambar 4. 9 Perbedaan Teks antara Kolom Ulasan dan <i>Kolom Tokenized</i>	93
Gambar 4. 10 Hasil dari Proses Normalisasi	95
Gambar 4. 11 Kode Program Mengembalikan ke Bentuk Kalimat	95
Gambar 4. 12 Kode <i>Python</i> pada Tahapan Penerjemah.....	96
Gambar 4. 13 Hasil Penerjemahan ke Kolom <i>English_Review</i>	98
Gambar 4. 14 Hasil Proses Pelabelan	99
Gambar 4. 15 Hasil Penghapusan Kolom <i>English_Review</i>	104
Gambar 4. 16 Proses Penghapusan Stopwords pada Ulasan.....	104
Gambar 4. 17 Tahapan <i>Stemming</i> Menggunakan <i>Library</i> Sastrawi	106
Gambar 4. 18 Proses Pemisahan Label	108
Gambar 4. 19 <i>Word Cloud</i> Sentimen Positif.....	110

Gambar 4. 20 <i>Word Cloud</i> Sentimen Negatif	111
Gambar 4. 21 Kode Program untuk Proses TF-IDF	114
Gambar 4. 22 Distribusi Sentimen Data Ulasan iPhone <i>Second</i> di Tokopedia ..	119
Gambar 4. 23 Proses Teknik <i>SMOTE</i>	120
Gambar 4. 24 Evaluasi Kinerja pada Model <i>Naïve Bayes</i>	133
Gambar 4. 25 <i>Confusion Matrix</i> dengan <i>Naïve Bayes</i>	134
Gambar 4. 26 Evaluasi Kinerja pada Model <i>Support Vector Machine</i>	137
Gambar 4. 27 <i>Confusion Matrix</i> dengan <i>Support Vector Machine</i>	138
Gambar 4. 28 Evaluasi kinerja model pada Algoritma SVM	141
Gambar 4. 29 Evaluasi kinerja model pada Algoritma <i>Naïve Bayes</i>	141
Gambar 4. 30 Tampilan fitur Uji Model SVM pada Hasil Prediksi	143
Gambar 4. 31 Tampilan fitur Uji Model SVM pada Nilai Akurasi	144
Gambar 4. 32 Tampilan fitur Uji Model SVM pada <i>Confusion Matrix</i>	145
Gambar 4. 33 Tampilan fitur Uji Model <i>Naïve Bayes</i> pada Hasil Prediksi	146
Gambar 4. 34 Tampilan fitur Uji Model <i>Naïve Bayes</i> pada Nilai Akurasi	147
Gambar 4. 35 Tampilan fitur Uji Model <i>Naïve Bayes</i> pada <i>Confusion Matrix</i> ..	148
Gambar 4. 36 Tampilan dari fitur Visualisasi pada aplikasi <i>web</i>	150
Gambar 4. 37 Tampilan dari fitur Visualisasi pada aplikasi <i>web</i>	150
Gambar 4. 38 Tampilan dari fitur Visualisasi pada aplikasi <i>web</i>	151
Gambar 4. 39 Tampilan dari fitur Prediksi Manual pada aplikasi <i>web</i>	152
Gambar 4. 40 Tampilan dari fitur Tentang pada aplikasi <i>web</i>	153