

**ANALISIS SENTIMEN ULASAN PRODUK *SNEAKERS*
LOKAL PADA TOKOPEDIA MENGGUNAKAN
ALGORITMA *NAIVE BAYES* DAN
*SUPPORT VECTOR MACHINE***

SKRIPSI

Oleh:

I Komang Arya Trisumeikra

202110715017



**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 *Sneakers*
Lokal Pada Tokopedia Menggunakan Algoritma
Naïve Bayes dan *Support Vector Machine*

Nama Mahasiswa : I Komang Arya Trisumeikra

Nomor Pokok Mahasiswa : 202110715017

Program Studi/Fakultas : Informatika/Ilmu Komputer

Tanggal Lulus Ujian Tugas Akhir : 25 Juli 2025

Jakarta, 03 Juli 2025

MENYETUJUI,

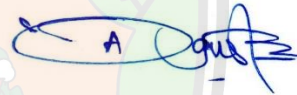
Pembimbing I

Pembimbing II



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

NIDN: 0311097302



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 *Sneakers*
Lokal Pada Tokopedia Menggunakan Algoritma
Naïve Bayes dan *Support Vector Machine*

Nama Mahasiswa : I Komang Arya Trisumeikra

Nomor Pokok Mahasiswa : 202110715017

Program Studi/Fakultas : Informatika/Illmu Komputer

Tanggal Lulus Ujian Tugas Akhir : 25 Juli 2025

Jakarta, 1 Agustus 2025

MENGESAHKAN,

Ketua Tim Penguji : Wowon Priatna, S.T., M.TI.

NIDN: 0429118007

Penguji I : Rafika Sari, S.Si., M.Si.

NIDN: 0329098902

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

NIDN: 0311097302

MENGETAHUI,

Ketua

Dekan

Program Studi Informatika

Fakultas Ilmu Komputer

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

NIP. 2012486

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 : I Komang Arya Trisumeikra
NPM : 202110715017
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Sentimen Ulasan Produk Sneakers Lokal Pada Tokopedia Menggunakan Algoritma Naïve Bayes dan Support Vector Machine

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila di kemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

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

Jakarta, 1 Agustus 2025

Penulis



I Komang Arya Trisumeikra

ABSTRAK

I Komang Arya Trisumeikra. 202110715017. Analisis Sentimen Ulasan Produk *Sneakers* Lokal Pada Tokopedia Menggunakan Algoritma *Naïve Bayes* dan *Support Vector Machine*. Informatika. Ilmu Komputer. Universitas Bhayangkara Jakarta Raya. 2025.

Perkembangan teknologi digital yang sangat pesat telah mendorong terjadinya transformasi signifikan dalam berbagai sektor bisnis, termasuk industri *fashion*, khususnya pada produk sepatu jenis *sneakers*. Produk ini tidak hanya berfungsi sebagai pelengkap penampilan, tetapi juga merepresentasikan gaya hidup modern, sehingga menjadi tren global. *Brand* ternama dunia seperti Nike, Adidas, dan Puma mendominasi pasar dengan pendapatan yang tinggi. Namun, harga jual yang relatif mahal menjadi kendala tersendiri bagi sebagian besar konsumen di Indonesia. Kondisi ini menciptakan peluang strategis bagi pelaku usaha lokal untuk menghadirkan produk *sneakers* dengan kualitas yang kompetitif dan harga yang lebih terjangkau melalui *e-commerce* seperti Tokopedia, yang merupakan *e-commerce* dengan trafik kunjungan tertinggi kedua di Indonesia. Penelitian ini bertujuan untuk menganalisis sentimen dari ulasan konsumen terhadap produk *sneakers* lokal guna memahami persepsi kualitas dan kepuasan produk serta mendukung pengambilan keputusan calon pembeli. Data diperoleh dari 1.032 ulasan pengguna di lima toko, yaitu NAH Project, Aerostreet, Geoff Max Official, Ventela_NEW, dan Brodo Footwear_NEW. Klasifikasi sentimen dilakukan menggunakan algoritma *Naïve Bayes* dan *Support Vector Machine* (SVM). Hasil evaluasi menunjukkan bahwa SVM memberikan akurasi tertinggi sebesar 98%, mengungguli *Naïve Bayes* dengan akurasi 96%. Model terbaik ini diimplementasikan dalam bentuk aplikasi berbasis web untuk analisis sentimen terhadap ulasan baru yang tersedia di kelima toko tersebut, khususnya untuk menilai kualitas produk *sneakers* lokal di Tokopedia.

Kata Kunci: Analisis Sentimen, *Sneakers* Lokal, Tokopedia, *Naïve Bayes*, *Support Vector Machine*.

ABSTRACT

I Komang Arya Trisumeikra. 202110715017. Sentiment Analysis of Local Sneakers Product Reviews on Tokopedia Using Naïve Bayes and Support Vector Machine Algorithm. Informatics. Computer Science. University of Bhayangkara Jakarta Raya. 2025.

The rapid advancement of digital technology has driven significant transformations across various business sectors, including the fashion industry, particularly in sneakers products. Sneakers today serve not only as fashion accessories but also as representations of a modern lifestyle, making them a global trend. Leading global brands such as Nike, Adidas, and Puma dominate the market with high revenues. However, their relatively high prices pose a challenge for many consumers in Indonesia. This situation presents a strategic opportunity for local businesses to offer competitively priced sneakers with comparable quality through e-commerce platforms such as Tokopedia, which ranks as the second most visited e-commerce site in Indonesia. This study aims to analysis consumer sentiment from reviews of local sneaker products to better understand consumer perceptions of product quality and satisfaction, thereby supporting prospective buyers in making informed decisions. The data were collected from 1,032 user reviews across five stores: NAH Project, Aerostreet, Geoff Max Official, Ventela_NEW, and Brodo Footwear_NEW. Sentiment classification was conducted using two machine learning algorithms: Naïve Bayes and Support Vector Machine (SVM). Evaluation results show that SVM achieved the highest accuracy at 98%, outperforming Naïve Bayes, which achieved 96%. The best-performing model was implemented in a web-based application to analysis sentiment in new reviews from the selected stores, particularly to assess the quality of local sneaker products on Tokopedia.

Keywords: Sentiment Analysis, Sneakers Local, Tokopedia, Naïve Bayes, Support Vector Machine.

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

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

Nama : I Komang Arya Trisumeikra
NPM : 202110715017
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 *Sneakers* Lokal Pada Tokopedia Menggunakan Algoritma *Naïve 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 izin 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 : Jakarta
Pada tanggal : 1 Agustus 2025
Yang Menyatakan



I Komang Arya Trisumeikra

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa, atas segala nikmat dan bimbingan nya yang telah memberikan kemudahan dan kelancaran, sehingga penulis dapat menyelesaikan tugas akhir skripsi yang berjudul **“Analisis Sentimen Terhadap Ulasan Produk *Sneakers* Lokal Pada Tokopedia Menggunakan Algoritma *Naïve Bayes* dan *Support Vector Machine*”**. Sebagai salah satu syarat yang harus dipenuhi dalam memperoleh gelar Sarjana (S1) Program Studi Informatika Fakultas Ilmu Komputer di Universitas Bhayangkara Jakarta Raya.

Dalam penyusunan tugas akhir skripsi ini penulis banyak mendapatkan bantuan dari berbagai pihak yang telah memberikan dukungan, oleh sebab itu penulis ingin mengucapkan rasa terima kasih kepada:

1. Bapak Irjen Pol. (Purn) Prof. 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, SE., M.M.S.I., selaku Ketua Program Studi Informatika Universitas Bhayangkara Jakarta Raya.
4. Ibu Dr. Herlawati, S.Si., M.M., M. Kom., selaku Dosen Pembimbing I yang telah membimbing, serta memberikan masukan dan motivasi selama penulis melakukan penyusunan penelitian ini.
5. Bapak Agus Hidayat, S. Kom., M.Si., selaku Dosen Pembimbing II yang telah membimbing, serta memberikan masukan dan motivasi selama penulis melakukan penyusunan penelitian ini.

6. Bapak Mayadi, S. Kom., M. Kom., selaku Dosen Pembimbing Akademik penulis selama menjadi mahasiswa di Universitas Bhayangkara Jakarta Raya.
7. Seluruh Dosen Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya yang telah mendidik dan memberikan ilmu selama kuliah dan seluruh jajaran staf atas bantuan dalam urusan pelayanan segala administrasi.
8. Kedua Orang Tua yang selalu memberi semangat, dukungan, dan mendoakan dengan penuh ketulusan dalam menyelesaikan penyusunan penelitian ini.
9. Para teman-teman, dan sahabat yang selalu memberikan dukungan dan membantu memberikan semangat selama penyusunan penelitian ini.

Penulis menyadari bahwa tugas akhir skripsi ini masih memiliki banyak kekurangan dan masih jauh dari sempurna. Oleh karena itu, segala masukan baik berupa kritik maupun saran yang membangun demi perbaikan kedepannya. Harapan penulis semoga tugas akhir skripsi ini bermanfaat untuk orang lain yang membacanya.

Jakarta, 1 Agustus 2025

Hormat saya



I Komang Arya Trisumeikra

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING.....	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xv
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	8
1.3 Rumusan Masalah	8
1.4 Batasan Masalah.....	9
1.5 Tujuan Penelitian.....	10
1.6 Manfaat Penelitian.....	10
1.7 Sistematika Penulisan.....	12
BAB II TINJAUAN PUSTAKA.....	14
2.1 Penelitian Terkait.....	14
2.2 Teori Pendukung.....	20
2.2.1 <i>Machine Learning</i>	20

2.2.2	<i>Data Mining</i>	21
2.2.3	<i>Cross-Industry Standard Process for Data mining (CRISP-DM)</i> ...	23
2.2.4	<i>Text Mining</i>	25
2.2.5	<i>Naïve Bayes</i>	27
2.2.6	<i>Support Vector Machine</i>	30
2.2.7	<i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	32
2.2.8	<i>Confusion Matrix</i>	34
2.2.9	<i>Synthetic Minority Over-sampling Technique (SMOTE)</i>	36
2.2.10	<i>Lexicon Based</i>	37
2.2.11	<i>Word Cloud</i>	37
2.2.12	<i>Analisis Sentimen</i>	38
2.2.13	<i>Web Scraping</i>	38
2.2.14	<i>Selenium</i>	39
2.2.15	<i>Beautifulsoup</i>	40
2.2.16	<i>Python</i>	40
2.2.17	<i>Sneakers</i>	41
2.2.18	<i>Tokopedia</i>	42
2.2.19	<i>Google Colab (Colaboratory)</i>	43
2.2.20	<i>Visual Studio Code</i>	44
BAB III METODOLOGI PENELITIAN		46
3.1	<i>Objek Penelitian</i>	46
3.2	<i>Kerangka Penelitian</i>	49
3.2.1	<i>Pemahaman Bisnis (Business Understanding)</i>	51
3.2.2	<i>Pemahaman Data (Data Understanding)</i>	51
3.2.3	<i>Persiapan Data (Data Preparation)</i>	54
3.2.4	<i>Pemodelan (Modeling)</i>	57

3.2.5	Evaluasi (<i>Evaluation</i>).....	61
3.2.6	Penerapan (<i>Deployment</i>)	61
3.3	Metode Pengumpulan Data	61
3.3.1	Observasi.....	62
3.3.2	Studi Pustaka.....	62
3.4	Analisis Kebutuhan Sistem	62
3.4.1	Perangkat Keras (<i>Hardware</i>)	63
3.4.2	Perangkat Lunak (<i>Software</i>).....	63
BAB IV HASIL DAN PEMBAHASAN.....		65
4.1	Pemahaman Bisnis (<i>Business Understanding</i>).....	65
4.2	Pemahaman Data (<i>Data Understanding</i>)	66
4.3	Persiapan Data (<i>Data Preparation</i>)	72
4.3.1	<i>Case Folding</i>	73
4.3.2	<i>Tokenizing</i>	74
4.3.3	<i>Normalize</i>	75
4.3.4	<i>Stopword Removal</i>	76
4.3.5	<i>Stemming</i>	77
4.3.6	Pelabelan	77
4.3.7	Visualisasi	79
4.3.8	<i>Splitting Data</i>	83
4.3.9	TF-IDF	85
4.3.10	SMOTE	90
4.4	Pemodelan (<i>Modeling</i>)	93
4.4.1	<i>Naïve Bayes</i>	93
4.4.2	<i>Support Vector Machine</i>	100
4.5	Evaluasi (<i>Evaluation</i>)	105

4.5.1	<i>Naïve Bayes</i>	106
4.5.2	<i>Support Vector Machine</i>	110
4.6	Penerapan (<i>Deployment</i>)	119
BAB V PENUTUP		126
5.1	Kesimpulan.....	126
5.2	Saran.....	127
DAFTAR PUSTAKA		130
LAMPIRAN		138



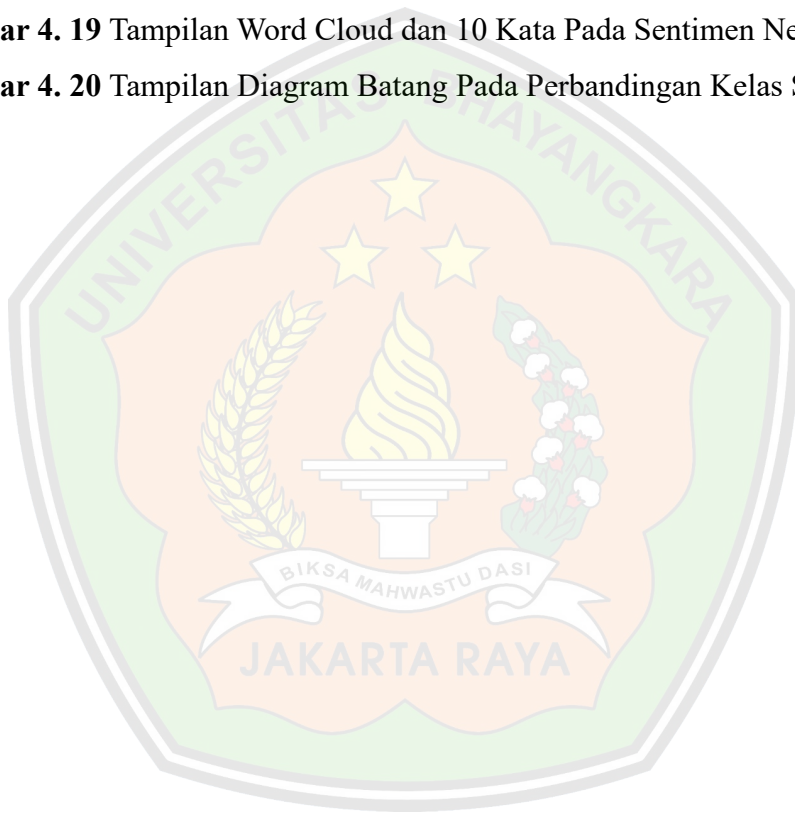
DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait	14
Tabel 3. 1 Spesifikasi Perangkat Keras	63
Tabel 3. 2 Spesifikasi Perangkat Lunak	64
Tabel 4. 1 Tahapan <i>Scraping</i>	66
Tabel 4. 2 Hasil Penerapan Tahapan <i>Case Folding</i>	74
Tabel 4. 3 Hasil Penerapan Tahap <i>Tokenizing</i>	74
Tabel 4. 4 Hasil Penerapan Tahap <i>Normalize</i>	75
Tabel 4. 5 Hasil Penerapan Tahap <i>Stopword Removal</i>	76
Tabel 4. 6 Hasil Penerapan Tahap <i>Stemming</i>	77
Tabel 4. 7 Hasil Jumlah Label Sentimen	79
Tabel 4. 8 Jumlah Kemunculan Kata Sentimen Positif	81
Tabel 4. 9 Jumlah Kemunculan Kata Sentimen Negatif	82
Tabel 4. 10 Distribusi Jumlah Data Latih dan Data Uji	84
Tabel 4. 11 Jumlah Data Transformasi Data Latih dan Data Uji	86
Tabel 4. 12 Sebagian Data Pada Perhitungan TF-IDF	86
Tabel 4. 13 Hasil Perhitungan TF Normalisasi	87
Tabel 4. 14 Hasil Perhitungan DF (<i>Document Frequency</i>)	88
Tabel 4. 15 Hasil Perhitungan IDF (<i>Inverse Document Frequency</i>)	89
Tabel 4. 16 Hasil Perhitungan TF-IDF	90
Tabel 4. 17 Distribusi Label Sebelum dan Sesudah SMOTE	91
Tabel 4. 18 Perhitungan Manual Pada <i>Data Training</i>	94
Tabel 4. 19 Perhitungan Manual Pada <i>Data Testing</i>	98
Tabel 4. 20 Hasil Probabilitas <i>Data Testing</i>	99
Tabel 4. 21 Hasil TF-IDF dari Tiga Ulasan	101
Tabel 4. 22 Hasil Perbandingan Evaluasi Pada <i>Naïve Bayes</i> dan SVM	114
Tabel 4. 23 Evaluasi Sebelum dan Sesudah SMOTE Pada <i>Naive Bayes</i>	115
Tabel 4. 24 Evaluasi Sebelum dan Sesudah SMOTE Pada SVM	116

DAFTAR GAMBAR

Gambar 1. 1	Data Pendapatan Sektor Sepatu <i>Brand</i> Nike, Adidas, dan Puma.....	2
Gambar 1. 2	Situs <i>E-commerce</i> dengan Pengunjung Terbanyak di Indonesia	4
Gambar 2. 1	Pengenalan Bagian Utama Pada <i>Machine Learning</i>	20
Gambar 2. 2	Tahapan dari CRISP-DM.....	23
Gambar 2. 3	<i>Confusion Matrix</i>	34
Gambar 2. 4	Penerapan Aplikasi <i>Python</i>	40
Gambar 2. 5	<i>Sneakers</i> Lokal	41
Gambar 2. 6	Logo Tokopedia	42
Gambar 2. 7	Logo <i>Google Colaboratory</i>	43
Gambar 2. 8	Logo <i>Visual Studio Code</i>	44
Gambar 3. 1	Halaman Ulasan Toko NAH Project di Tokopedia.....	47
Gambar 3. 2	Halaman Ulasan Toko Aerostreet di Tokopedia	47
Gambar 3. 3	Halaman Ulasan Toko Geoff Max Official di Tokopedia.....	48
Gambar 3. 4	Halaman Ulasan Toko Ventela_NEW di Tokopedia.....	48
Gambar 3. 5	Halaman Ulasan Toko Brodo Footwear_NEW di Tokopedia.....	49
Gambar 3. 6	Kerangka Penelitian.....	50
Gambar 3. 7	Hasil <i>Scraping</i> Ulasan pada Toko NAH Project.....	52
Gambar 3. 8	Hasil Penggabungan Dataset	53
Gambar 3. 9	Ulasan Produk yang Bukan <i>Sneakers</i>	53
Gambar 3. 10	Tahapan Pemodelan Algoritma <i>Naïve Bayes</i>	57
Gambar 3. 11	Tahapan Pemodelan Algoritma <i>Support Vector Machine</i>	59
Gambar 4. 1	Kode Penggabungan Dataset	70
Gambar 4. 2	Hasil Penggabungan Dataset	71
Gambar 4. 3	Kode Filter Dataset.....	71
Gambar 4. 4	Hasil Filter Dataset	72
Gambar 4. 5	Hasil Penghapusan Duplikasi Dataset	73
Gambar 4. 6	Hasil Penerapan Tahap Pelabelan	78
Gambar 4. 7	<i>Word Cloud</i> Sentimen Positif.....	80
Gambar 4. 8	<i>Word Cloud</i> Sentimen Negatif.....	82
Gambar 4. 9	Grafik Distribusi Label Sebelum dan Sesudah SMOTE	92

Gambar 4. 10 Hasil Evaluasi <i>Naïve Bayes</i>	106
Gambar 4. 11 <i>Confusion Matrix Naive Bayes</i>	107
Gambar 4. 12 Hasil Evaluasi <i>Support Vector Machine</i>	110
Gambar 4. 13 <i>Confusion Matrix Support Vector Machine</i>	111
Gambar 4. 14 Tampilan Antarmuka Fitur Prediksi Manual	120
Gambar 4. 15 Tampilan Implementasi Pada Kalimat Ulasan Positif	121
Gambar 4. 16 Tampilan Implementasi Pada Kalimat Ulasan Negatif.....	121
Gambar 4. 17 Tampilan Antarmuka Fitur Visualisasi	122
Gambar 4. 18 Tampilan <i>Word Cloud</i> dan 10 Kata Pada Sentimen Positif	123
Gambar 4. 19 Tampilan Word Cloud dan 10 Kata Pada Sentimen Negatif.....	123
Gambar 4. 20 Tampilan Diagram Batang Pada Perbandingan Kelas Sentimen .	124



DAFTAR LAMPIRAN

Lampiran 1 Cek Plagiarisme

Lampiran 2 Biodata Mahasiswa

Lampiran 3 Kartu Konsultasi Bimbingan Pembimbing I

Lampiran 4 Kartu Konsultasi Bimbingan Pembimbing II

Lampiran 5 Surat Rekomendasi Pembimbing

Lampiran 6 Dataset Sekunder Ulasan Produk *Sneakers* Lokal

