

**TUGAS AKHIR**  
**ANALISIS SENTIMEN KOMENTAR PRODUK**  
**KECANTIKAN XYZ DI APLIKASI *E-COMMERCE***  
***SHOPEE* MENGGUNAKAN METODE KLASIFIKASI**  
***NAÏVE BAYES***



Oleh :

Veronika Cahya Putri

1462000026

**PROGRAM STUDI TEKNIK INFORMATIKA**  
**FAKULTAS TEKNIK**  
**UNIVERSITAS 17 AGUSTUS 1945 SURABAYA**  
**2024**

*Halaman ini sengaja dikosongkan*

**TUGAS AKHIR**

**ANALISIS SENTIMEN KOMENTAR PRODUK  
KECANTIKAN XYZ DI APLIKASI *E-COMMERCE*  
*SHOPEE* MENGGUNAKAN METODE KLASIFIKASI  
*NAÏVE BAYES***

**Diajukan sebagai salah satu syarat untuk memperoleh gelar  
Sarjana Komputer di Program Studi Informatika**



**Oleh :**

**Veronika Cahya Putri**

**1462000026**

**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS 17 AGUSTUS 1945 SURABAYA  
2024**

*Halaman ini sengaja dikosongkan*

**FINAL PROJECT**  
**SENTIMEN ANALYSIS OF XYZ BEAUTY PRODUCT**  
**COMMENTS ON THE SHOPEE E-COMMERCE**  
**APPLICATION USING THE *NAÏVE BAYES***  
***CLASSIFICATION METHOD***

**Prepared as partial fulfilment of the requirement for the degree of  
Sarjana Komputer at Informatics Departement**



**By :**  
**Veronika Cahya Putri**  
**1462000026**

**INFORMATICS DEPARTEMENT**  
**FACULTY OF ENGINEERING**  
**UNIVERSITAS 17 AGUSTUS 1945 SURABAYA**  
**2024**

*Halaman ini sengaja dikosongkan*

**PROGRAM STUDI INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS 17 AGUSTUS 1945 SURABAYA**

---

**LEMBAR PENGESAHAN TUGAS AKHIR**

**Nama** : Veronika Cahya Putri  
**NBI** : 1462000026  
**Prodi** : S-1 Informatika  
**Fakultas** : Teknik  
**Judul** : ANALISIS SENTIMEN KOMENTAR PRODUK  
KECANTIKAN XYZ DI APLIKASI E-COMMERCE  
SHOPEE MENGGUNAKAN METODE KLASIFIKASI  
NAÏVE BAYES

**Mengetahui / Menyetujui**  
**Dosen Pembimbing**



Andrey Kartika Widhy Hapantenda, S.Kom., M.Kom.  
NPP. 20460.19.0804

**Dekan Fakultas Teknik**

**Universitas 17 Agustus 1945  
Surabaya**



Dr. Ir. Sajyo, ST., M.Kes., IPU., ASEAN Eng  
NPP. 20410.90.0197

**Ketua Program Studi Informatika**

**Universitas 17 Agustus 1945  
Surabaya**



Aidil Primasetya Armin, S.ST., M.MT.  
NPP. 20460.16.0700

*Halaman ini sengaja dikosongkan*

# PERNYATAAN KEASLIAN DAN PERSETUJUAN PUBLIKASI TUGAS AKHIR

Saya yang bertanda tangan di bawah ini:

Nama : Veronika Cahya Putri  
NBI : 1462000026  
Fakultas/Program Studi : Teknik/Informatika  
Tugas Akhir : Analisis Sentimen Komentar Produk Kecantikan XYZ  
Di Aplikasi E-Commerce Shopee Menggunakan  
Metode Klasifikasi Naïve Bayes

Menyatakan dengan sesungguhnya bahwa:

1. Tugas Akhir dengan judul di atas bukan merupakan tiruan atau duplikasi dari Tugas Akhir yang sudah dipublikasikan dan atau pernah dipakai untuk mendapatkan gelar Sarjana Teknik di lingkungan Universitas 17 Agustus 1945 Surabaya maupun di Perguruan Tinggi atau Instansi manapun, kecuali bagian yang sumber informasinya dicantumkan sebagaimana mestinya.
2. Tugas Akhir dengan judul diatas bukan merupakan plagiarisme, pencurian hasil karya milik orang lain, hasil kerja orang lain untuk kepentingan saya karena hubungan material maupun non-material, ataupun segala kemungkinan lain yang pada hakekatnya bukan merupakan karya tulis tugas akhir saya secara orisinal dan otentik.
3. Demi pengembangan ilmu pengetahuan, saya memberikan hak atas Tugas Akhir ini kepada Universitas 17 Agustus 1945 Surabaya untuk menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.
4. Pernyataan ini saya buat dengan kesadaran sendiri dan tidak atas tekanan ataupun paksaan dari pihak maupun demi menegakan integritas akademik di institusi ini dan bila kemudian hari diduga kuat ada ketidaksesuaian antara fakta dengan kenyataan ini, saya bersedia diproses oleh tim Fakultas yang dibentuk untuk melakukan verifikasi, dengan sanksi terberat berupa pembatalan kelulusan/kesarjanaan.

Surabaya, 11 Juli 2024



Veronika Cahya Putri  
1462000026

*Halaman ini sengaja dikosongkan*



UNIVERSITAS  
17 AGUSTUS 1945  
SURABAYA

**BADAN  
PERPUSTAKAAN**

JL. SEMOLOWARU 45 SURABAYA TLP.  
031 593 1800 (EX 311)  
EMAIL: [PERPUS@UNTAG-SBY.AC.ID](mailto:PERPUS@UNTAG-SBY.AC.ID)

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA  
ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas 17 Agustus 1945 Surabaya, Saya yang bertanda tangan di bawah ini:

Nama : Veronika Cahya Putri  
NIM : 1462000026  
Fakultas : Teknik  
Program Studi : Informatika  
Jenis Karya : Tugas Akhir/Skripsi

Demi perkembangan ilmu pengetahuan, saya meyetujui untuk memberikan kepada Badan Perpustakaan Universitas 17 Agustus 1945 Surabaya **Hak Bebas Royalti Noneklusif (Nonexclusive Royalty-Free Right)**, atas karya saya yang berjudul:

**Analisis Sentimen Komentar Produk Kecantikan XYZ Di Aplikasi E-Commerce Shopee Menggunakan Metode Klasifikasi Naïve Bayes**

Dengan **Hak Bebas Royalti Noneklusif (Nonexclusive Royalty-Free Right)**, Badan Perpustakaan Universitas 17 Agustus 1945 Surabaya berhak menyimpan, mengalihkan media atau memformatkan, mengolah dalam bentuk pangkalan data (database), merawat, mempublikasikan karya ilmiah saya selama tetap tercantum.

Dibuat di : Universitas 17 Agustus 1945 Surabaya  
Pada Tanggal : 11 Juli 2024

Yang Menyatakan



Veronika Cahya Putri)

*Halaman ini sengaja dikosongkan*

## KATA PENGANTAR

Puji syukur kepada Allah Yang Maha Esa dan Yang Maha Kuasa yang senantiasa telah melimpahkan Rahmat dan Hidayah-Nya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul “ANALISIS SENTIMEN KOMENTAR PRODUK KECANTIKAN XYZ DI APLIKASI E-COMMERCE SHOPEE MENGGUNAKAN ALGORITMA NAIVE BAYES” sebagai salah satu persyaratan untuk menyelesaikan studi di Universitas 17 Agustus 1945 Surabaya dan mendapatkan gelar Sarjana. Saya menyadari bahwa bantuan dari Allah dan doa orang tua, serta dukungan dari teman-teman seangkatan, sangatlah penting bagi penulis sehingga dapat menyelesaikan Tugas Akhir dengan baik dan tepat waktu.

Selain itu, penulis juga ingin menyampaikan ucapan terima kasih yang mendalam kepada pihak-pihak berikut:

1. Bapak Andrey Kartika Widhy Hapantenda, S.Kom., M.Kom. selaku dosen pembimbing saya, yang telah memberikan petunjuk, pengarahan, semangat dan bimbingan dari awal pembuatan Tugas Akhir dan yang telah memberikan banyak waktunya untuk *sharing* berbagai hal sehingga penulis dapat menyelesaikan Tugas Akhir ini.
2. Bapak Dosen Wali Mohamad Sidqon.S.,Kom.,M.Si yang telah membimbing dan mengarahkan saya selama masa studi di Universitas 17 Agustus Surabaya ini.
3. Bapak Aidil Primasetya Armin, S.ST., M.T. selaku Ketua Program Studi Teknik Informatika Universitas 17 Agustus 1945 Surabaya.
4. Keluarga tercinta, Bapak Ibu sebagai orang tua, yang selalu mendoakan, memotivasi, memperhatikan, dan melengkapkan segala keperluan penulis hingga terselesaikannya Tugas Akhir ini.
5. Teman-teman satu angkatan dan satu perjuangan yang telah melewati proses Tugas Akhir bersama.
6. M.Fachri Chusaini yang senantiasa mendampingi serta membantu saya dalam menyelesaikan Tugas Akhir ini.
7. Semua pihak yang tidak dapat penulis sebutkan satu per satu yang telah membantu dalam bentuk apapun dalam penyelesaian tugas akhir ini.

*Halaman ini sengaja dikosongkan*

# DAFTAR ISI

<b>LEMBAR PENGESAHAN TUGAS AKHIR .....</b>	<b>i</b>
<b>PERNYATAAN KEASLIAN DAN PERSETUJUAN PUBLIKASI TUGAS AKHIR... ..</b>	<b>iii</b>
<b>KATA PENGANTAR.....</b>	<b>v</b>
<b>DAFTAR ISI.....</b>	<b>vii</b>
<b>DAFTAR GAMBAR.....</b>	<b>ix</b>
<b>DAFTAR TABEL.....</b>	<b>xi</b>
<b>ABSTRAK .....</b>	<b>xiii</b>
<b>ABSTRACT .....</b>	<b>xv</b>
<b>BAB 1 PENDAHULUAN .....</b>	<b>1</b>
1.1. Latar Belakang.....	1
1.2. Perumusan Masalah .....	2
1.3. Batasan Masalah .....	2
1.4. Tujuan Penelitian .....	2
1.5. Manfaat Penelitian .....	3
<b>BAB 2 TINJAUAN PUSTAKA DAN DASAR TEORI .....</b>	<b>5</b>
2.1. Penelitian Terdahulu .....	5
2.2. Dasar Teori .....	7
2.2.1. Analisis Sentimen.....	7
2.2.2. <i>Naïve Bayes</i> .....	9
2.2.3. Corpus .....	11
2.2.4. <i>Natural Language Processing (NLP)</i> .....	12
2.2.5. <i>Web Scraping</i> .....	13
<b>BAB 3 METODOLOGI PENELITIAN .....</b>	<b>15</b>
3.1. Bahan dan Perangkat Penelitian .....	15
3.1.1. Perangkat Keras (Hardware) .....	15
3.1.2. Perangkat Lunak (Software).....	15
3.2. Objek Penelitian.....	15
3.3. Tahapan Penelitian.....	15
3.3.1. Studi Literatur .....	16
3.3.2. Pengumpulan Data ( <i>Web Scraping</i> ) .....	16
3.3.3. Tahap Pra-pemrosesan .....	18
3.3.4. Tahap Pelabelan Kelas .....	18
3.3.5. Tahap Pemisahan Data Training, Data Validasi & Data Uji.....	19
3.3.6. Implementasi Data.....	20
3.3.7. Evaluasi .....	20
3.4. Skenario Pengujian .....	20
3.5. Perancangan Sistem .....	20

3.5.1.	Use Case Diagram .....	21
3.5.2.	Activity Diagram .....	22
3.5.3.	Sequence Diagram .....	23
3.6.	Perancangan desain aplikasi (Mockup ).....	23
3.6.1.	Rancangan Halaman Utama .....	24
3.6.2.	Rancangan Halaman Input Link .....	25
3.6.3.	Rancangan Halaman Loading Indicator .....	25
3.6.4.	Rancangan Halaman Tampilan Tabel Ringkas.....	26
3.6.5.	Rancangan Halaman Tampilan Tabel Lengkap.....	27
3.6.6.	Rancangan Halaman Visualisasi Data .....	27
<b>BAB 4</b>	<b>HASIL DAN PEMBAHASAN .....</b>	<b>29</b>
4.1.	Tahap pengumpulan data .....	29
4.2.	<i>Text Processing</i> .....	29
4.3.	Labeling .....	33
4.4.	Hasil dan Evaluasi.....	40
4.5.	Implementasi Sistem .....	57
<b>BAB 5</b>	<b>PENUTUP .....</b>	<b>61</b>
5.1.	Kesimpulan .....	61
5.2.	Saran.....	61
<b>DAFTAR PUSTAKA</b>	<b>.....</b>	<b>63</b>

## DAFTAR GAMBAR

<b>Gambar 2.1.</b> Analisis Sentimen pada Review Produk.....	7
<b>Gambar 2.2.</b> Teknik Analisis Sentimen untuk Klasifikasi.....	8
<b>Gambar 2.3.</b> NLP dengan Machine Learning.....	12
<b>Gambar 2.4.</b> Web Scraping.....	13
<b>Gambar 2.5.</b> Web Scraping Process .....	14
<b>Gambar 3.1.</b> Alur Penelitian .....	16
<b>Gambar 3.2.</b> Alur Kerja Web Scraping.....	17
<b>Gambar 3.3.</b> Tahap Pra-pemrosesan dengan NLP.....	18
<b>Gambar 3.4.</b> Use Case Diagram .....	21
<b>Gambar 3.5.</b> Activity Diagram .....	22
<b>Gambar 3.6.</b> Sequence Diagram .....	23
<b>Gambar 3.7.</b> Halaman Utama .....	24
<b>Gambar 3.8</b> Input Link .....	25
<b>Gambar 3.9</b> Proses Scraping.....	25
<b>Gambar 3.10</b> Halaman Tampilan Ringkas.....	26
<b>Gambar 3.11</b> Halaman Tampilan Tabel Lengkap.....	27
<b>Gambar 3.12</b> Halaman Tampilan Visualisasi .....	28
<b>Gambar 4.1.</b> Alur Text Processing .....	30
<b>Gambar 4.2</b> Hasil Tokenizing dan Postag .....	31
<b>Gambar 4.3</b> Hasil perhitungan bobot kata .....	34
<b>Gambar 4.4</b> Wordcloud Positif.....	34
<b>Gambar 4.5</b> Hasil Visualisasi data sampel satu .....	35
<b>Gambar 4.6</b> Hasil Oversampling Negatif 1500 dan Netral 700.....	36
<b>Gambar 4.7.</b> Hasil Oversampling Netral 1000 dan Negatif 1000.....	36
<b>Gambar 4.8.</b> Hasil Oversampling Negatif 500 dan Netral 1000.....	37
<b>Gambar 4.9</b> Hasil Oversampling Negatif 1500 dan Netral 1500.....	38
<b>Gambar 4.10</b> Hasil Oversampling Negatif 2000 dan Netral 2000.....	39
<b>Gambar 4.11</b> Hasil Oversampling dengan data yang sama 2893 .....	39
<b>Gambar 4.12</b> Tampilan Halaman Awal Aplikasi.....	57
<b>Gambar 4.13</b> Tampilan Input Link .....	57
<b>Gambar 4.14.</b> Tampilan Loading Proses Scraping dan Processing .....	58
<b>Gambar 4.15.</b> Tampilan Hasil Scraping dan Processing Data. ....	58
<b>Gambar 4.16.</b> Tampilan Hasil Scraping dan Processing Data Lengkap .....	59
<b>Gambar 4.17.</b> Hasil dan Visualisasi.....	60

*Halaman ini sengaja dikosongkan*

## DAFTAR TABEL

<b>Tabel 3.1.</b> Contoh Pelabelan Kelas (Kata) .....	19
<b>Tabel 3.2.</b> Pembagian data latih, data validasi, dan data uji .....	19
<b>Tabel 3.3.</b> Use Case Sistem .....	21
<b>Tabel 4.1.</b> Contoh Hasil scraping .....	29
<b>Tabel 4.2</b> Kamus Kata Alay .....	31
<b>Tabel 4.3.</b> Contoh Hasil Stopword .....	32
<b>Tabel 4.4.</b> Tabel Steaming.....	32
<b>Tabel 4.5</b> Lexicon negatif.....	33
<b>Tabel 4.6</b> Lexicon Positif.....	33
<b>Tabel 4.7.</b> Contoh Prediksi Naive Bayes Berdasarkan Probabilitas .....	41
<b>Tabel 4.8.</b> Data original Proporsi 60 : 40 .....	42
<b>Tabel 4.9.</b> Data asli Proporsi 70:30 .....	43
<b>Tabel 4.10</b> Data Oversampling 1500 Negatif 700 Netral Proporsi 70:30 .....	44
<b>Tabel 4.11</b> Data Oversampling 1500 Negatif 700 Netral Proporsi 60:40 .....	45
<b>Tabel 4.12</b> Data Oversampling 1000 Negatif 1000 Netral Proporsi 70:30 .....	46
<b>Tabel 4.13</b> Data Oversampling 1000 Negatif 1000 Netral Proporsi 60:40 .....	47
<b>Tabel 4.14.</b> Data Oversampling 500 Negatif 1000 Netral Proporsi 70:30 .....	48
<b>Tabel 4.15</b> DataOversampling Negatif 500 dan Netral 1000 proporsi 60:40.....	49
<b>Tabel 4.16</b> Data Oversampling 1500 Netral dan 1500 Negatif Proporsi 70:30.....	50
<b>Tabel 4.17</b> Data Oversampling Negatif 1500 dan Netral 1500 Proporsi 60:40.....	51
<b>Tabel 4.18</b> Data Oversampling Netral 2000 dan Negatif 2000 proporsi 70:30 .....	52
<b>Tabel 4.19</b> Data Oversampling Negatif 2000 dan Netral 2000 proporsi 60 :40.....	53
<b>Tabel 4.20</b> Data Oversampling 2893 Netral dan 2893 Negatif proporsi 70:30 .....	54
<b>Tabel 4.21</b> Data Oversampling 2893 Negatif dan 2893 Netral proporsi 60:40 .....	55

*Halaman ini sengaja dikosongkan*

## ABSTRAK

Nama : Veronika Cahya Putri  
Program Studi : Teknik Informatika  
Judul : Analisis Sentimen Komentar Produk Kecantikan XYZ Di  
Aplikasi *E-Commerce Shopee* Menggunakan Metode  
Klasifikasi *Naïve Bayes*

Di era modern ini, belanja online melalui aplikasi e-commerce semakin populer di Indonesia. Salah satu produk yang sering dibeli adalah produk kecantikan, dengan 80% masyarakat Indonesia berbelanja secara daring dan 63,3% dari mereka menggunakan aplikasi e-commerce Shopee. Shopee menjadi e-commerce dengan jumlah pengunjung terbanyak di Indonesia pada kuartal pertama 2023. Ulasan atau komentar pada produk di e-commerce dapat membantu menentukan kesesuaian produk dengan harapan konsumen. Untuk menentukan klasifikasi label pada ulasan produk (negatif, positif, dan netral), digunakan metode Naive Bayes. Data untuk pengujian diperoleh melalui teknik web scraping. Namun, data yang didapatkan tidak seimbang, sehingga perlu dilakukan teknik oversampling untuk menyeimbangkan distribusi kelas. Setelah melalui proses text processing dan oversampling, hasil pengujian menunjukkan akurasi model Naive Bayes mencapai 98% dengan proporsi data 70:30 (pelatihan:pengujian). Penggunaan teknik oversampling pada data yang tidak seimbang terbukti dapat meningkatkan kinerja model Naive Bayes dalam mengklasifikasikan ulasan produk kecantikan pada aplikasi e-commerce.

**Kata Kunci:** Aplikasi *E-Commerce Shopee*, *Naïve Bayes*, Produk kecantikan, Ulasan atau Komentar, *Web Scraping*.

*Halaman ini sengaja dikosongkan*

## ABSTRACT

Name : Veronika Cahya Putri  
Department : Informatics  
Title : *Sentimen Analysis Of XYZ Beauty Product Comments On The Shopee E-Commerce Application Using The Naïve Bayes Classification Method*

In the modern era, online shopping through e-commerce applications is increasingly popular in Indonesia. One of the frequently purchased products is beauty products, with 80% of Indonesians shopping online and 63.3% of them using the Shopee e-commerce app. Shopee has become the most visited e-commerce platform in Indonesia as of the first quarter of 2023. Reviews or comments on products in e-commerce can help determine the product's suitability to consumer expectations. To classify product review labels (negative, positive, and neutral), the Naive Bayes method is employed. Testing data is obtained through web scraping, but it is unbalanced, necessitating oversampling techniques to balance class distributions. After undergoing text processing and oversampling, the testing results show that the Naive Bayes model achieves 98% accuracy with a data proportion of 70:30 (training ). The use of oversampling on unbalanced data proves effective in enhancing the performance of the Naive Bayes model in classifying beauty product reviews on e-commerce applications.

**Keywords:** Application Shopee, Beauty Products, E-Commerce, Reviews or Comments Naïve Bayes, Web Scraping.

*Halaman ini sengaja dikosongkan*