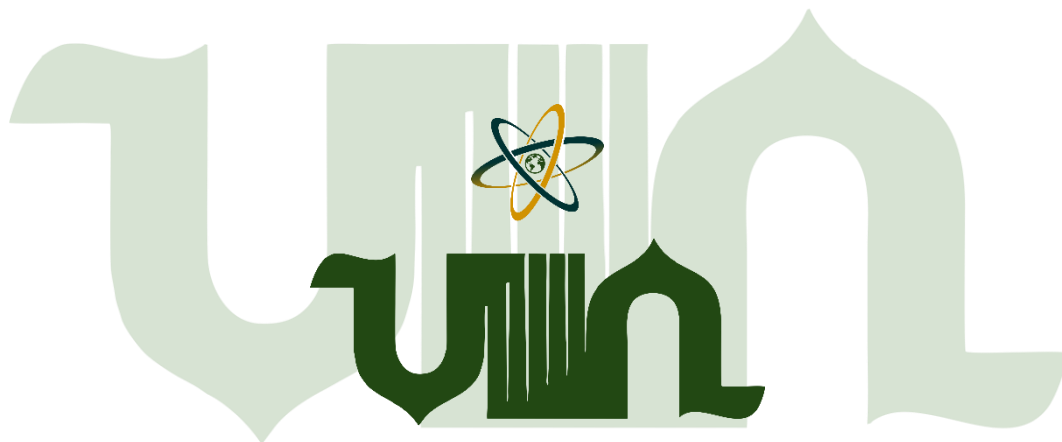


**ANALISIS SENTIMEN PADA *REVIEW* PENGGUNA
E-COMMERCE MENGGUNAKAN ALGORITMA NAÏVE
BAYES (STUDI KASUS : SHOPEE)**

SKRIPSI

**DINDA ZUKHOIRIYAH
0701183278**



UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA
MEDAN**

2022

**ANALISIS SENTIMEN PADA *REVIEW* PENGGUNA
E-COMMERCE MENGGUNAKAN ALGORITMA NAÏVE
BAYES (STUDI KASUS : SHOPEE)**

SKRIPSI

Diajukan untuk Memenuhi Syarat Mencapai Gelar Sarjana Komputer

DINDA ZUKHOIRIYAH
0701183278



UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA
MEDAN**

2022

LEMBAR PESRSETUJUAN

Hal : Surat Persetujuan Skripsi

Lamp : -

Kepada Yth.,
Dekan Fakultas Sains dan Teknologi
Universitas Islam Negeri Sumatera Utara Medan

Assalamu'alaikum Wr. Wb

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengatakan perbaikan, maka kami selaku pembimbing berpendapat bahwa skripsi saudara,

Nama : Dinda Zukhoiriyah
Nomor Induk Mahasiswa : 0701183278
Program Studi : Ilmu Komputer
Judul : Analisis Sentimen Pada *Review* Pengguna *E-Commerce* Menggunakan Algoritma Naïve Bayes (Studi Kasus: Shopee).

Dapat disetujui untuk segera di *Munaqasyahkan*. Atas perhatiannya kami mengucapkan terimakasih.

Medan, 31 Agustus 2022

03 Safar 1444 H

Komisi Pembimbing,

Pembimbing Skripsi I,



Abdul Halim Hasugian, M.Kom
NIP. 1100000113

Pembimbing Skripsi II,



Dr. M. Fakhriza, M.Kom
NIB.1100000115

SURAT PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini;

Nama : Dinda Zukhoiriyah
Nomor Induk Mahasiswa : 071183278
Program Studi : Ilmu Komputer
Judul : Analisis Sentimen Pada *Review* Pengguna
E-Commerce Menggunakan Algoritma
Naïve Bayes (Studi Kasus : Shopee)

menyatakan bahwa skripsi ini adalah hasil karya saya sendiri, kecuali beberapa kutipan dari ringkasan yang masing-masing disebutkan sumbernya. Apabila di kemudian hari ditemukan plagiat dalam skripsi ini. Maka saya bersedia menerima sanksi pencabutan gelar akademik yang saya peroleh dan sanksi lainnya sesuai dengan peraturan yang berlaku.

Medan, 31 Agustus 2022



Dinda Zukhoiriyah
0701183278

UNIVERSITAS ISLAM SUMATERA UTARA
SUMATERA UTARA MEDAN



KEMENTERIAN AGAMA REPUBLIK INDONESIA
UNIVERSITAS ISLAM NEGERI SUMATERA UTARA MEDAN
FAKULTAS SAINS DAN TEKNOLOGI

Jln.Lapangan Golf, Desa Durian Jangak, Kecamatan Pancur Batu, Kabupaten
Deli Serdang

Provinsi Sumatera Utara, Kode Pos : 20353

Telp.(061) 6615683-6622925, Fax. (061) 6615683

URL: www.saintek.uinsu.ac.id, email: saintek@uinsu.ac.id

PENGESAHAN SKRIPSI

Nomor: B.265/ST/ST.V.2/PP.01.1/10/2022

Judul : Analisis Sentimen Pada *Review* Pengguna *E-Commerce* Menggunakan Algoritma Naïve Bayes (Studi Kasus : Shopee)
Nama : Dinda Zukhoiriyah
Nomor Induk Mahasiswa : 07011832782
Program Studi : Ilmu Komputer
Fakultas : Sains dan Teknologi

Telah dipertahankan dihadapan Dewan Penguji Skripsi Studi Ilmu Komputer Fakultas Sains dan Teknologi UIN Sumatera Utara Medan dan dinyatakan **LULUS**.

Pada hari/tanggal : Rabu, 31 Agustus 2022
Tempat/media : Ruang Sidang Fak. Sains dan Teknologi
UINSU Medan Kampus IV Tuntungan

Tim Ujian Munaqsyah,
Ketua

Ilka Zufria, M.Kom
NIP. 198506042015031006

Dewan Penguji,

Penguji I,

Ilka Zufria, M.Kom
NIP.198506042015031006

Penguji II,

Sulindawaty, M.Kom
NIP.0107048201

Penguji III,

Abdul Halim Hasugian, M.Kom
NIB.1100000113

Penguji IV,

Dr. M.Fakhriza, M.Kom
NIB.1100000115

Mengesahkan,
Dekan Fakultas Sains dan Teknologi
UIN Sumatera Utara



Prof. Dr. M. Syafran, M.A
NIP.196609051991031002

ABSTRAK

Ulasan produk merupakan salah satu sumber informasi tentang kualitas produk dan sangat berpengaruh pada konsumen dan produsen. Menghadapi jumlah data yang besar dan sangat banyak untuk setiap produk pada sebuah toko di Shopee, menganalisis dan menyimpulkan informasi secara manual membutuhkan waktu lama dan tidak efektif. Sistem analisis sentimen dibutuhkan untuk mengekstrak informasi penting dan membangun sistem yang dapat menentukan kualitas produk secara objektif dan menangani informasi teks yang besar. Sistem analisis sentimen memiliki beberapa tahap yaitu *crawling*, *preprocessing*, pembobotan kata, dan klasifikasi sentimen. Penulis menggunakan *API* untuk mengumpulkan data dan melakukan simulasi dengan menggunakan 600 data review pada website Shopee untuk menghasilkan sentimen positif dan negatif. Dataset dibagi menjadi dua yaitu data latih dan data uji dengan pembagian 400 data latih dan 200 data uji dari total dataset. Klasifikasi sentimen data latih dilakukan dengan menggunakan metode Lexicon Based. Pada pelabelan ini dilakukan secara otomatis dengan metode menghitung skor nilai sentimen memakai kamus Lexicon. Penelitian ini menggunakan klasifikasi dengan Algoritma Naïve Bayes menggunakan pengujian *confusion matrix* untuk mendapatkan hasil *accuracy*, *precision* dan *recall*. Hasil pengujian menunjukkan nilai *accuracy* sebesar 99,5%, *precision* sebesar 99,49%, *recall* sebesar 100%. Dengan demikian dapat disimpulkan metode klasifikasi algoritma Naïve Bayes cukup relevan meskipun akurasi nya belum 100%.

Kata Kunci : Analisis Sentimen, Lexicon Based, Naïve Bayes, *Review*, Shopee.

UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

ABSTRACT

Product reviews are a source of information about product quality and are very influential to consumers and producers. Faced with large and overwhelming amounts of data for each product in a shop on Shopee, analyzing and summarizing information manually takes a long time and is ineffective. A sentiment analysis system is needed to extract important information and build a system that can objectively determine product quality and handle large textual information. The sentiment analysis system has several stages, namely crawling, preprocessing, word weighting, and sentiment classification. The author uses the API to collect data and perform simulations using 600 data reviews on the Shopee website to generate positive and negative sentiments. The dataset is divided into two, namely training data and test data by dividing 400 training data and 200 test data from the total dataset. The training data sentiment classification is carried out using the Lexicon Based method. This labeling is done automatically by calculating the sentiment value score using the Lexicon dictionary. This study uses classification with the Naïve Bayes Algorithm using confusion matrix testing to get accuracy, precision and recall results. The test results show an accuracy value of 99.5%, a precision of 99.49%, a recall of 100%. Thus it can be concluded that the classification method of the Naïve Bayes algorithm is quite relevant even though its accuracy is not 100%.

Keywords : Sentiment Analysis, Lexicon Based, Naïve Bayes, Review, Shopee.



UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

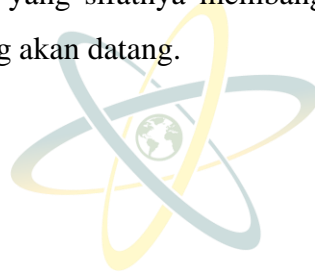
KATA PENGANTAR

Puji Syukur penulis ucapkan kepada Allah SWT yang telah memberikan rahmat beserta hidayah-Nya sehingga penulis dapat menyelesaikan proposal skripsi yang berjudul “Analisis Sentimen Pada *Review* Pengguna *E-Commerce* Menggunakan Algoritma Naïve Bayes (Studi Kasus : Shopee)”.

Penulis sangat berterima kasih kepada semua pihak yang telah membantu dan mendukung sehingga proposal skripsi ini bisa diselesaikan. Penulis ingin mengucapkan terima kasih kepada :

1. Orangtua Penulis, Ayah yang tak kenal lelah memberikan dukungan semangat dan motivasi sehingga penulis dapat menyelesaikan skripsi ini.
2. Bapak Prof. Dr. Syahrin Harahap, MA selaku Rektor Universitas Islam Negeri Sumatera Utara
3. Bapak Prof. Dr. Mhd. Syahnan, MA selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sumatera Utara
4. Bapak Ilka Zufria, M.Kom selaku Ketua Program Studi Ilmu Komputer Universitas Islam Negeri Sumatera Utara
5. Bapak Abdul Halim Hasugian, M.Kom selaku Dosen Pembimbing I yang telah memberikan arahan, nasihat dan informasi sehingga saya dapat menyelesaikan skripsi ini.
6. Bapak Dr. M. Fakhriza, M.Kom selaku Dosen Pembimbing II yang telah memberikan arahan, nasihat dan informasi sehingga saya dapat menyelesaikan skripsi ini.
7. Bapak Dr. Mhd. Furqan, S.Si.,M.Comp.Sc selaku Dosen Pembimbing Akademik yang telah membimbing dan memberikan arahan dalam pengajuan judul skripsi.
8. Seluruh Dosen serta staff pegawai di Program Studi Ilmu Komputer Fakultas Sains dan Teknologi Universitas Islam Negeri Sumatera Utara
9. Seluruh keluarga penulis yang sangat peduli dan memberikan dukungannya.
10. Teman-teman stambuk 2018, khususnya kelas ILKOMP 6 yang selama ini telah memberikan banyak pelajaran hidup.

11. Abang dan kakak senior yang telah memberikan masukan yang telah banyak membantu dan memberikan semangat.
12. Semua pihak yang terlibat langsung ataupun tidak langsung yang tidak dapat diucapkan satu-persatu yang telah membantu penyelesaian skripsi ini. Semoga Allah SWT memberikan berkah, kesehatan dan keselamatan bagi semua pihak yang telah mendukung penulis untuk menyelesaikan skripsi ini. Penulis berharap semoga skripsi ini dapat bermanfaat dan menambah wawasan keilmuan. Kritik dan saran yang sifatnya membangun sangat penulis harapkan untuk perbaikan dimasa yang akan datang.



Medan, Agustus 2022

Penulis,

Dinda Zukhoiriyah

UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

DAFTAR ISI

	Halaman
ABSTRAK	i
ABSTRACT	ii
KATA PENGANTAR	iii
DAFTAR ISI	v
DAFTAR GAMBAR	vii
DAFTAR TABEL	ix
DAFTAR LAMPIRAN	xi
BAB I PENDAHULUAN	1
1.1.Latar Belakang	1
1.2.Rumusan Masalah	4
1.3.Batasan Masalah.....	4
1.4.Tujuan Penelitian	4
1.5.Manfaat Penelitian	5
BAB II TINJAUAN PUSTAKA	6
2.1.Analisis Sentimen.....	6
2.2.E-Commerce	6
2.3.Shopee	7
2.4.API Shopee.....	8
2.5.Text Mining.....	8
2.6.Preprocessing Data.....	9
2.7.Lexicon Based.....	13
2.8.Fitur Term Frequency	14
2.9.Klasifikasi	15
2.10. Kriteria Komentar Positif dan Negatif	15
2.11.Algoritma Naïve Bayes.....	16
2.12.Teknik Klasifikasi Naïve Bayes.....	16
2.13.Confussion Matrix	22
2.14.Flowchart	23
2.15.HTML (<i>HyperText Markup Language</i>).....	25

2.16.PHP (Personal Home Page)	25
2.17.Laragon	27
2.18.MYSQL.....	28
2.19.Database.....	29
2.20.Sublime Text.....	30
2.21.Penelitian Terdahulu	30
BAB III METODOLOGI PENELITIAN	35
3.1.Waktu dan Jadwal Penelitian	35
3.2.Bahan dan Alat Penelitian.....	35
3.2.1. Perangkat Keras	35
3.2.2. Perangkat Lunak.....	36
3.3.Prosedur Kerja.....	36
3.3.1. Tahap Persiapan	36
3.3.2. Teknik Pengumpulan Data.....	37
3.3.3. Analisi Kebutuhan.....	38
3.3.4. Tahap Perancangan	38
BAB IV HASIL DAN PEMBAHASAN.....	41
4.1. Pembahasan.....	41
4.1.1. Analisis Data.....	42
4.1.2. Representasi Data.....	44
4.1.2.1. Preprocessing Data.....	49
4.1.3. Hasil Analisis Data	54
4.1.3.1. Pembobotan.....	54
4.1.3.2. Klasifikasi Naïve Bayes.....	56
4.2. Hasil	64
4.2.1. Pengujian.....	74
4.2.2. Penerapan.....	74
BAB V KESIMPULAN DAN SARAN	75
5.1. Kesimpulan	75
5.2. Saran.....	75
DAFTAR PUSTAKA	76
LAMPIRAN-LAMPIRAN	

DAFTAR GAMBAR

Gambar	Judul Gambar	Halaman
2.1	<i>Block Diagram Cleaning</i>	9
2.2	Contoh <i>Cleaning</i>	9
2.3	Alur Proses <i>Case Folding</i>	10
2.4	Contoh <i>Case Folding</i>	10
2.5	Alur Proses Normalisasi.....	11
2.6.	Contoh Normalisasi.....	11
2.7.	Contoh <i>Stemming</i>	12
2.8.	Contoh <i>Filtering</i>	12
2.9.	Contoh <i>Tokenizing</i>	13
2.10.	Alur Proses Pelabelan dengan <i>Lexicon Based</i>	14
3.1.	Kerangka Kerja Penelitian	37
3.2.	<i>Flowchart</i> Sistem	39
4.1	Alur Sistem <i>Crawling</i> data <i>Review</i> Produk Shopee.....	41
4.2	Tampilan menu <i>get ratings</i> Pada API Shopee	42
4.3	Tampilan Tahap <i>Crawling</i> Data pada Aplikasi Sublime Text..	42
4.4	Data Hasil <i>Crawling</i>	43
4.5	Block Diagram <i>Preprocessing</i> Data.....	48
4.6	Menu Beranda.....	64
4.7	Halaman Menu <i>Login</i>	65
4.8	Halaman Utama <i>Dashboard</i>	65
4.9	Halaman Menu Cari Toko.....	66
4.10	Menu Halaman Produk	66
4.11	Menu Info Produk	67
4.12	Menu Analisis Sentimen	68
4.13	Tampilan Menu Data Latih	68
4.14	Tampilan Visualisasi Hasil Analisis Sentimen Positif.....	69
4.15	Tampilan Visualisasi Hasil Analisis Sentimen Negatif	70
4.16	Tampilan Visualisasi Hasil Analisis Data Latih Positif.....	71

4.17 Tampilan Visualisasi Hasil Analisis Data Latih Negatif	71
4.18 Tampilan Visualisasi Hasil Analisis Data Uji.....	72



UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

DAFTAR TABEL

Tabel	Judul Tabel	Halaman
2.1	Contoh Term Sentimen Positif.....	18
2.2	Contoh Term Sentimen Negatif	19
2.3	Contoh Term Sentimen Netral	19
2.4	Contoh Perubahan Probabilitas Sentimen.....	20
2.5	Contoh Probabilitas <i>Term</i> Data Uji Pada Sentimen Positif.....	21
2.6	Probabilitas Term Data Uji Pada Sentimen Negatif	21
2.7	Probabilitas Term Data Uji Pada Sentimen Netral	22
2.8	Confusion Matrix	22
2.9	Simbol-simbol Flowchart.....	23
2.10	Penelitian Terdahulu	30
3.1	Waktu dan Jadwal Pelaksanaan Penelitian	35
4.1	Kata Terdeteksi Oleh Kamus Liu Lexicon.....	44
4.2	Data Latih.....	44
4.3	Data Uji	46
4.4	Sampel Data Latih.....	47
4.5	Sampel Data Uji	48
4.6	Proses <i>Cleaning</i>	49
4.7	Proses <i>Case Folding</i>	50
4.8	Proses Normalisasi	50
4.9	Proses <i>Filtering</i>	51
4.10	Proses <i>Stemming</i>	52
4.11	Proses <i>Tokenizing</i>	53
4.12	Frekuensi Data Latih Positif.....	54
4.13	Frekuensi Data Latih Negatif	55
4.14	Frekuensi Data Uji	55
4.15	Perubahan Probabilitas Term Pada Sentimen Positif.....	59
4.16	Perubahan Probabilitas Term Pada Sentimen Negatif	60
4.17	Perbandingan Jumlah Kelas Sentimen Data Latih	72

4.18 Perbandingan Jumlah Kelas Sentimen Data Uji	72
4.19 Perbandingan Jumlah Kelas Sentimen Data Actual.....	73
4.20 Confusion Matrix Data Uji.....	73
4.21 Hasil Evaluasi.....	74



UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN

DAFTAR LAMPIRAN

Lampiran	Judul Lampiran
1.	Data Latih
2.	Data Uji
3.	Listing Program
4.	Daftar Riwayat Hidup
5.	Kartu Bimbingan Skripsi



UNIVERSITAS ISLAM NEGERI
SUMATERA UTARA MEDAN