

***SENTIMENT ANALYSIS ULASAN PRODUK PADA SITUS E-
MARKETPLACE TOKOPEDIA DENGAN MENGGUNAKAN TEXT MINING***

SKRIPSI

**Diajukan untuk memenuhi sebagian persyaratan
dalam memperoleh gelar Sarjana Komputer
Program Studi Sistem Informasi
Disusun Oleh:**



RAHMANDA WAHYU ADINEGORO

1635010028

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAWA TIMUR
S U R A B A Y A**

2020

SKRIPSI

SENTIMENT ANALYSIS* ULASAN PRODUK PADA SITUS *E-MARKETPLACE* TOKOPEDIA DENGAN MENGGUNAKAN *TEXT MINING

Disusun Oleh:

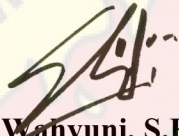
RAHMANDA WAHYU ADINEGORO

1635010028

Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Skripsi
Program Studi Sistem Informasi Fakultas Ilmu Komputer
Universitas Pembangunan Nasional “Veteran” Jawa Timur
Pada Tanggal 6 Juli 2020

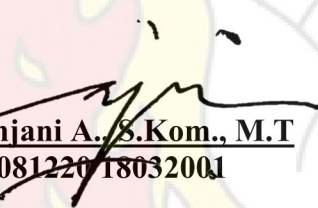
Dosen Pembimbing :

1.



Eka Dyar Wahyuni, S.Kom, M.Kom
NPT. 3 8412 13 0356 1

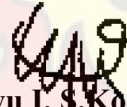
2.



Amalia Anjani A., S.Kom., M.T
NIP. 199208122018032001

Dosen Penguji :

1.




Syurfah Ayu I., S.Kom, M.Kom
NPT. 3 8501 10 0294 1

2.



Asif Faroqi, S.Kom., M.Kom
NIP. 1987051920 18031001

3.



Eristya Maya Safitri, S.Kom, M.Kom.
NIP. 1993031620 19032020

Mengetahui,

Dekan fakultas ilmu komputer

Universitas Pembangunan “Veteran” Jawa Timur



Dr. Ir. Ni Ketut Sari, M.T

NIP. 19650731 199203 2 001

SURAT PERNYATAAN

Saya, Mahasiswa Sistem Informasi UPN “Veteran” Jawa Timur, yang bertanda tangan di bawah ini :

Nama : Rahmanda Wahyu Adinegoro

NPM : 1635010028

Program Studi : Sistem Informasi

Fakultas : Ilmu Komputer

Menyatakan Bahwa Judul Skripsi / Tugas Akhir Saya Sebagai Berikut :

SENTIMENT ANALYSIS ULASAN PRODUK PADA SITUS E-MARKETPLACE TOKOPEDIA DENGAN MENGGUNAKAN TEXT MINING

Bukan merupakan plagiat dari skripsi / tugas akhir / penelitian orang lain dan juga bukan merupakan produk / software / hasil karya yang saya beli dari pihak lain.

Saya juga menyatakan bahwa Skripsi / Tugas Akhir ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di UPN “Veteran” Jawa Timur maupun institusi pendidikan lain.

Jika ternyata dikemudian hari pernyataan ini terbukti tidak benar, maka saya bertanggung jawab penuh dan siap menerima segala konsekuensinya, termasuk pembatalan ijazah dikemudian hari.



Hormat Saya,

Rahmanda Wahyu Adinegoro

Judul : ***SENTIMENT ANALYSIS* ULASAN PRODUK
PADA SITUS E-MARKETPLACE TOKOPEDIA
DENGAN MENGGUNAKAN TEXT MINING**

Penulis : **Rahmanda Wahyu Adinegoro**

Pembimbing 1 : **Eka Dyar Wahyuni, S.Kom., M.Kom.**

Pembimbing 2 : **Amalia Anjani Arifiyanti, S.Kom., M.Kom**

ABSTRAK

Indonesia memiliki penduduk dengan penggunaan internet yang cukup besar, pada tahun 2018 tercatat jika pengguna internet di Indonesia mencapai 64,8% dari total penduduk sebanyak 264,16 juta jiwa. Hal ini mengakibatkan peningkatan transaksi melalui e-commerce, dimana Indonesia menyumbang sebesar 52,3% nilai transaksi e-commerce yang terjadi di Asia Tenggara pada tahun 2018. Salah satu situs e-marketplace terbesar di Indonesia adalah Tokopedia, pada situs ini terdapat ulasan dan rating produk. Tetapi dalam penerapannya banyak ulasan yang tak sesuai dengan rating produk yang diberikan, dan banyak penjual yang memaksa pembeli untuk memberikan rating yang tinggi, maka dari permasalahan ini dapat dilakukan analisa sentimen yang terkandung dalam ulasan yang diberikan pengguna.

Sentimen yang dianalisa merupakan sentimen positif, netral, dan negatif. Untuk mendapatkan ulasan produk pada Tokopedia dilakukan pengambilan data dengan cara mengakses API Tokopedia, lalu dilakukan text mining untuk mengetahui sentimen yang terkandung dalam suatu ulasan. Text mining dilakukan dengan algoritma Naive Bayes menggunakan 4 algoritma naive bayes yang berbeda untuk mengetahui algoritma dengan performa terbaik.

Dari penelitian yang dilakukan, algoritma Naive Bayes dengan performa terbaik adalah Complement dimana menghasilkan tingkat akurasi prediksi sebesar 83,66%, dengan akurasi sebesar 84,37% pada pengujian sistem pada data ulasan sebanyak 32 buah.

Kata kunci : *Analysis Sentimen, Tokopedia, Naive Bayes, Klasifikasi, Data Mining*

KATA PENGANTAR

Puji dan syukur saya panjatkan kehadirat Allah SWT, atas segala nikmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul “SENTIMENT ANALYSIS ULASAN PRODUK PADA SITUS E-MARKETPLACE TOKOPEDIA DENGAN MENGGUNAKAN TEXT MINING”. Skripsi ini disusun dalam rangka memenuhi persyaratan dalam menyelesaikan Program Studi S1 / Sistem Informasi di Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Dalam penyusunan skripsi ini saya menyadari jika skripsi ini tidak akan terwujud tanpa bantuan dan dukungan dari berbagai pihak. Oleh sebab itu saya ingin mengucapkan terimakasih kepada :

1. Bapak Rachmad Hidayat dan Ibu Indah Dwi Sari selaku orang tua saya yang selalu memberikan do'a, serta motivasi untuk kelancaran skripsi ini.
2. Ibu Eka Dyar Wahyuni, S.Kom., M.Kom., selaku dosen pembimbing 1 serta dosen wali yang senantiasa memberikan bimbingan dan saran terkait penyelesaian skripsi ini.
3. Ibu Amalia Anjani Arifiyanti, S.Kom., M.Kom, selaku dosen pembimbing 2 yang senantiasa memberikan bimbingan dan arahan dalam penyelesaian skripsi ini.
4. Seluruh teman-teman Sistem Informasi angkatan 2016 yang selalu kompak dan saling mendukung.

Saya menyadari bahwa penulisan skripsi ini masih jauh dari kata sempurna. Saya berharap skripsi ini dapat menunjang perkembangan ilmu pengetahuan, khususnya dalam bidang Sistem Informasi.

Surabaya,6 Juli 2020

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN	ii
KETERANGAN REVISI	iii
SURAT PERNYATAAN	iv
ABSTRAK	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xii
BAB I	1
1.1 LATAR BELAKANG.....	1
1.2 RUMUSAN MASALAH	3
1.3 BATASAN MASALAH	3
1.4 TUJUAN	3
1.5 SISTEMATIKA PENULISAN	4
BAB II.....	5
2.1 Penelitian Terdahulu.....	5
2.2 Dasar Teori	6
2.2.1 E-Marketplace	6
2.2.2 Tokopedia.....	7
2.2.3 Python	7
2.2.4 Django.....	7
2.2.5 <i>Sentiment Analysis</i>	8
2.2.6 <i>Term Frequency – Inverse Document Frequency</i>	8
2.2.7 <i>Naïve Bayes</i>	9
BAB III	10
3.1 Studi Literatur.....	11
3.2 Analisis Kebutuhan	11

3.2.1	Kebutuhan Data Penelitian.....	11
3.2.2	Kebutuhan Hardware dan Software	11
3.3	Pembangunan Model Klasifikasi.....	11
3.3.1	Pengumpulan Data	13
3.3.2	Pelabelan Data.....	13
3.3.3	Preprocess Data.....	13
3.3.4	Pembobotan TF-IDF	16
3.3.5	Pembagian Data	16
3.3.6	Klasifikasi Data.....	17
3.4	Evaluasi Performa Model Klasifikasi.....	17
3.5	Perancangan Sistem.....	18
3.5.1	Diagram Alir Aplikasi Web Analisis Sentimen Tokopedia.....	19
3.5.2	Perancangan Antarmuka	19
3.5.3	Perancangan <i>Data Flow Diagram</i> (DFD).....	20
3.6	Implementasi Sistem	20
3.7	Pengujian Sistem	20
BAB IV	21
4.1	Analisis Kebutuhan	21
4.1.1	Kebutuhan Data.....	21
4.1.2	Kebutuhan Software dan Hardware	21
4.2	Pembangunan Model Klasifikasi.....	22
4.2.1.	Pengumpulan Data	22
4.2.2.	Pelabelan Data.....	26
4.2.3.	Preprocessing Data	27
4.2.4.	Pembobotan TF-IDF	33
4.2.5.	Pembagian Data	34
4.2.6.	Klasifikasi Data.....	34

4.2.7.	Evaluasi Performa Model Klasifikasi	35
4.3	Perancangan Sistem.....	46
4.3.1	Perancangan Antarmuka Aplikasi Web	46
4.3.2	Perancangan <i>Data Flow Diagram</i>	49
4.4	Implementasi Sistem	50
4.5	Pengujian Sistem	57
BAB V	64
5.1	Kesimpulan.....	64
5.2	Saran	64
REFERENSI	666

DAFTAR GAMBAR

Gambar 1.1 Diagram Batang Pertumbuhan Transaksi E-Commerce (Google & TEMASEK,2018)	1
Gambar 3.1 Diagram Alir Penelitian	10
Gambar 3.2 Diagram Alir Pembangunan Model	12
Gambar 3.3 Diagram Alir Pengumpulan Data.....	13
Gambar 3.4 Diagram Alir Preprocessing Data	14
Gambar 3.5 Diagram Alir Aplikasi Web Analisis Sentimen Tokopedia.....	19
Gambar 4.1 Import Library yang dibutuhkan untuk pengambilan data.....	22
Gambar 4.2 Source Code Pengambilan Cookies	23
Gambar 4.3 Source Code Pengambilan Data (1)	24
Gambar 4.4 Source Code Pengambilan Data (2)	25
Gambar 4.5 Source Code untuk merubah ulasan kedalam bentuk (.csv).....	26
Gambar 4.6 Source Code Case Folding	27
Gambar 4.7 Source Code Cleansing Data.....	28
Gambar 4.8 Source Code Stopword Removal	29
Gambar 4.9 Source Code Proses Tokenize	30
Gambar 4.10 Source Code Proses Spell Correction (1).....	31
Gambar 4.11 Source Code Proses Spell Correction (2).....	31
Gambar 4.12 Source Code Proses Stemming	32
Gambar 4.13 Source Code Proses Pembobotan TF-IDF	33
Gambar 4.14 Contoh Hasil Proses Pembobotan TF-IDF.....	34
Gambar 4.15 Source Code Proses Pembagian Data.....	34
Gambar 4.16 Source Code Multinomial Naive Bayes	35
Gambar 4.17 Source Code Gaussian Naive Bayes	35
Gambar 4.18 Source Code Complement Naive Bayes.....	35
Gambar 4.19 Source Code Bernoulli Naive Bayes.....	35
Gambar 4.20 Source Code Evaluasi Performa Model Klasifikasi Algoritma Multinomial Naive Bayes	36

Gambar 4.21 Source Code Evaluasi Performa Model Klasifikasi Algoritma Gaussian Naïve Bayes	36
Gambar 4.22 Source Code Evaluasi Performa Model Klasifikasi Algoritma Complement Naïve Bayes.....	36
Gambar 4.23 Source Code Evaluasi Performa Model Klasifikasi Algoritma Bernoulli Naïve Bayes	36
Gambar 4.24 Perancangan Halaman Index.....	46
Gambar 4.25 Perancangan Halaman Analisa.....	47
Gambar 4.26 Perancangan Halaman Hasil.....	48
Gambar 4.27 Perancangan Halaman Tentang.....	49
Gambar 4.28 DFD Context Website Sentimen Analisis Ulasan Tokopedia.....	50
Gambar 4.29 DFD Level 0 Website Sentimen Analisis Ulasan Tokopedia	50
Gambar 4.30 Source Code import Library dan Load Data Training	51
Gambar 4.31 Source Code Pengambilan Cookies	52
Gambar 4.32 Implementasi Halaman Index.....	53
Gambar 4.33 Implementasi Halaman Analisis.....	53
Gambar 4.34 Implementasi Source Code Sentiment Analysis (1).....	54
Gambar 4.35 Implementasi Source Code Sentiment Analysis (2).....	54
Gambar 4.37 Implementasi Halaman Hasil (1).....	55
Gambar 4.38 Implementasi Halaman Hasil (2).....	56

DAFTAR TABEL

Tabel 2.1 Tabel Penelitian Terdahulu	6
Tabel 4.1 Tabel Contoh Hasil Pengumpulan Data.....	26
Tabel 4.2 Hasil Pelabelan Data Ulasan	27
Tabel 4.3 Hasil Proses Case Folding	27
Tabel 4.4 Contoh Hasil Proses Data Cleansing	29
Tabel 4.5 Contoh Hasil Proses Stopword Removal	30
Tabel 4.6 Contoh Hasil Proses Tokenize	31
Tabel 4.7 Contoh Hasil Proses Spell Correction.....	32
Tabel 4.8 Contoh Hasil Proses Stemming	33
Tabel 4.9 Tabel Confusion Matrix Algoritma Multinomial Naive Bayes	37
Tabel 4.10 Performa Algoritma Multinomial Naive Bayes	37
Tabel 4.11 Tabel Confusion Matrix Algoritma Gaussian Naive Bayes.....	39
Tabel 4.12 Performa Algoritma Gaussian Naive Bayes	39
Tabel 4.13 Tabel Confusion Matrix Algoritma Complement Naive Bayes.....	41
Tabel 4.14 Performa Algoritma Complement Naive Bayes	41
Tabel 4.15 Tabel Confusion Matrix Algoritma Bernoulli Naive Bayes	43
Tabel 4.16 Performa Algoritma Bernoulli Naive Bayes.....	43
Tabel 4.17 Evaluasi Performa Algoritma Model Klasifikasi.....	44
Tabel 4.18 Hasil Pengujian Prediksi Sistem	61