

**ANALISIS SENTIMEN TERHADAP TINGKAT KEPUASAN PENGGUNA  
PENYEDIA LAYANAN TELEKOMUNIKASI SELULER DI INDONESIA  
PADA MEDIA SOSIAL TWITTER MENGGUNAN ALGORITMA NAIVE  
BAYES**

**SKRIPSI**



**disusun oleh**

**Yuli Kurniawati**

**15.11.9169**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2020**

**ANALISIS SENTIMEN TERHADAP TINGKAT KEPUASAN PENGGUNA  
PENYEDIA LAYANAN TELEKOMUNIKASI SELULER DI INDONESIA  
PADA MEDIA SOSIAL TWITTER MENGGUNAN ALGORITMA NAIVE  
BAYES**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Informatika



**disusun oleh**

**Yuli Kurniawati**

**15.11.9169**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2020**

# PERSETUJUAN

## SKRIPSI

### ANALISIS SENTIMEN TERHADAP TINGKAT KEPUASAN PENGGUNA PENYEDIA LAYANAN TELEKOMUNIKASI SELULER DI INDONESIA PADA MEDIA SOSIAL TWITTER MENGGUNAN ALGORITMA NAIVE BAYES

yang dipersiapkan dan disusun oleh

**Yuli Kurniawati**

**15.11.9169**

telah disetujui oleh Dosen Pembimbing Skripsi

**Dosen Pembimbing,**

**Dr. Kusrini, M.Kom.**

**NIK. 190302106**

**PENGESAHAN**

**SKRIPSI**

**ANALISIS SENTIMEN TERHADAP TINGKAT KEPUASAN  
PENGGUNA PENYEDIA LAYANAN TELEKOMUNIKASI SELULER DI  
INDONESIA PADA MEDIA SOSIAL TWITTER MENGGUNAN  
ALGORITMA NAIVE BAYES**

yang dipersiapkan dan disusun oleh

**Yuli Kurniawati**

**15.11.9169**

telah dipertahankan di depan Dewan Penguji  
pada tanggal 21 Juli 2020

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Dr. Kusriani, M.Kom.**  
**NIK. 190302106**

**Mardhiya Hayaty, S.T., M.Kom.**  
**NIK. 190302108**

**Ike Verawati, M.Kom.**  
**NIK. 190302237**

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer

**DEKAN FAKULTAS ILMU KOMPUTER**

**Krisnawati, S.Si, M.T.**

**NIK. 190302038**

## PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 21 Juli 2020



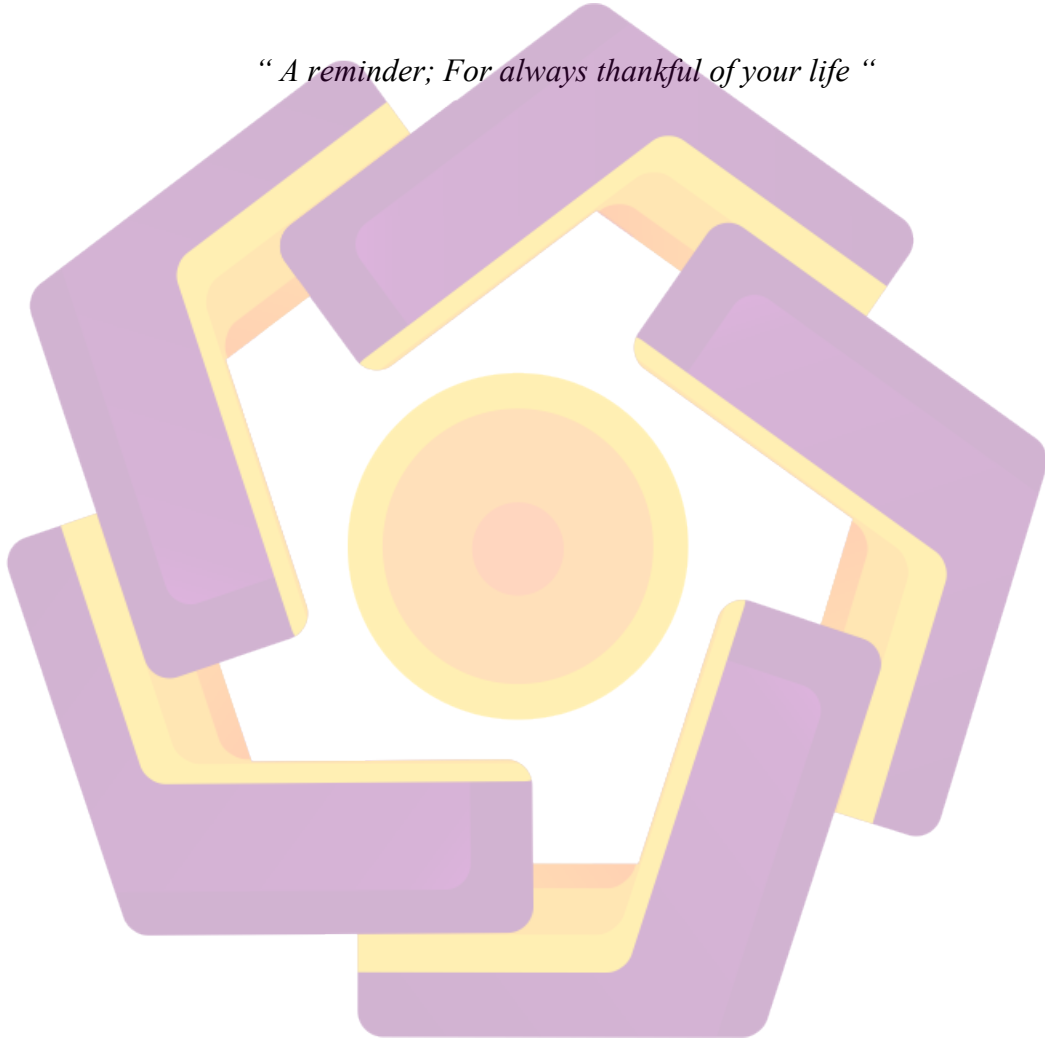
Yuli Kurniawati

NIM. 15.11.9169

## MOTTO

*“ There’s one happy thing happening everyday, even if you’re gloomy” (Jung Yong Hwa)*

*“ A reminder; For always thankful of your life “*



## PERSEMBAHAN

Alhamdulillah segala puji bagi Allah SWT. Tuhan Yang Maha Pengasih lagi Maha Penyayang, yang telah memberikan kemudahan bagi penulis menyelesaikan skripsi ini. Dengan segala ucapan terimakasih penulis persembahkan skripsi ini kepada orang-orang terkasih.

1. Allah SWT, atas rahmat, hidayah, serta karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik.
2. Kedua Orangtua tercinta, terimakasih atas segala dukungan dan doa yang tidak pernah berhenti. Alhamdulillah saya telah berhasil mewujudkan keinginan Bapak dan Ibu untuk menyelesaikan pendidikan di jenjang Sarjana, meskipun selesai terlambat.
3. Ibu Dr. Kusriani, M.Kom, selaku dosen pembimbing. Terimakasih atas arahan dan bimbingannya selama penulis mengerjakan skripsi ini.
4. Dian Indriastuti, Kharismah Nugroho, Shoffi Az-Zuhroh, Syifa Junisa, terimakasih telah menemani penulis selama masa kuliah.
5. Teman-teman yang tidak bisa penulis sebutkan satu per satu yang selalu mendorong penulis untuk menyelesaikan skripsi ini.
6. Tidak lupa ucapan terimakasih juga penulis persembahkan untuk yang selalu bertanya kapan skripsi selesai.

Untuk semua pihak yang telah penulis sebutkan mapupun yang tidak dapat disebutkan satu persatu, terimakasih banyak. Semoga Tuhan senantiasa membalas semua kebaikan kalian, serta semoga urusan kalian dimudahkan oleh Allah SWT.

## KATA PENGANTAR

Puji Syukur kehadiran Allah SWT yang telah memberikan rahmat dan hidayah-Nya kepada penulis sehingga dapat menyelesaikan Tugas Akhir dengan judul “Analisis Sentimen Tingkat Kepuasan Pengguna Penyedia Layanan Telekomunikasi Seluler di Indonesia pada Media Sosial Twitter Menggunakan Algoritma Naïve Bayes”. Skripsi ini disusun sebagai salah satu syarat kelulusan Program Sarjana Jurusan Informatika Universitas Amikom Yogyakarta dan memperoleh gelar Sarjana Komputer.

Dalam penyusunan tugas akhir ini, tentu masih banyak kekurangan dan hambatan yang terkadang ditemui sehingga dalam melengkapi penyusunan tugas akhir ini tidak lepas dari bantuan, dorongan, dan bimbingan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis memberikan ucapan terimakasih kepada :

1. Bapak Prof. Dr. M. Suyanto, MM, selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Krisnawati, S.Si, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Bapak Sudarmawan, M.T. selaku Ketua Program Studi Informatika.
4. Ibu Dr. Kusriani, M.Kom selaku dosen pembimbing yang telah memberikan pengarahan dan bimbingan kepada penulis.
5. Bapak dan Ibu dosen Universitas Amikom Yogyakarta yang telah membagi ilmunya kepada peneliti.



6. Kedua orangtua beserta keluarga yang selalu memberikan motivasi, doa, serta dorongan.
7. Teman-teman dan pihak lain yang selalu memberikan dukungan selama pengerjaan Skripsi ini.

Penulis menyadari bahwa dalam penyusunan Skripsi ini masih banyak kelemahan dan kekurangan. Oleh karena itu saran dan masukan yang membangun dari pembaca sangat diharapkan sebagai acuan untuk lebih baik lagi di waktu yang akan datang. Semoga Skripsi ini dapat bermanfaat bagi semua pihak yang membacanya.

Yogyakarta, 21 Juli 2020

Penulis

## DAFTAR ISI

PERSETUJUAN .....	iii
PENGESAHAN .....	iv
PERNYATAAN .....	v
MOTTO .....	vi
PERSEMBAHAN .....	vii
KATA PENGANTAR .....	viii
DAFTAR ISI .....	x
DAFTAR TABEL .....	xiii
DAFTAR GAMBAR .....	xiv
INTISARI .....	xv
<i>ABSTRACT</i> .....	xvi
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	2
1.3 Batasan Masalah .....	3
1.4 Maksud dan Tujuan Penelitian .....	3
1.5 Manfaat Penelitian .....	3
1.6 Metode Penelitian .....	4
1.6.1 Metode Pengumpulan Data .....	4
1.6.2 Metode Analisis .....	4
1.6.3 Metode Perancangan .....	4
1.6.4 Implementasi .....	5
1.6.5 Pengujian .....	5

1.7	Sistematika Penulisan.....	5
<b>BAB II LANDASAN TEORI.....</b>		<b>7</b>
2.1	Kajian Pustaka.....	7
2.2	Dasar Teori.....	9
2.2.1	Twitter dan Twitter API.....	9
2.2.2	Klasifikasi.....	10
2.2.3	<i>Text Mining</i> .....	11
2.2.4	<i>Data Mining</i> .....	11
2.2.5	<i>Preprocessing</i> .....	11
2.2.6	Naïve Bayes.....	13
2.2.7	Pembobotan Kata.....	15
2.2.8	Confusion Matrix.....	16
<b>BAB III ANALISIS DAN PERANCANGAN.....</b>		<b>19</b>
3.1	Analisis Masalah.....	19
3.2	Solusi yang Dipilih.....	20
3.3	Analisis Kebutuhan.....	20
3.3.1	Analisis Kebutuhan Fungsional.....	21
3.3.2	Analisis Kebutuhan Non Fungsional.....	21
3.4	Analisa Algoritma.....	21
3.4.1	Case Folding.....	21
3.4.2	Tokenizing.....	23
3.4.3	Filtering.....	24
3.4.4	Stemming.....	25
3.4.5	Klasifikasi Menggunakan Algoritma Naïve Bayes.....	27
3.5	Perancangan Aplikasi.....	38

3.5.1	Flowchart Sistem.....	38
3.5.2	Diagram Alur Preprocessing .....	39
3.5.3	Proses Klasifikasi .....	40
3.6	Perancangan Antar Muka .....	40
3.6.1	Antar Muka Halaman Awal .....	40
3.6.2	Antar Muka Halaman Hasil Klasifikasi .....	41
3.7	Perancangan Download Tweet.....	41
<b>BAB IV</b>	<b>IMPLEMENTASI DAN PEMBAHASAN .....</b>	<b>43</b>
4.1	Implementasi .....	43
4.1.1	Deskripsi Implementasi.....	43
4.1.2	Pengumpulan Data Tweet .....	43
4.1.3	Implementasi Antar Muka .....	45
4.2	Pembahasan Source Code .....	48
4.2.1	Source Code Pengambilan Data.....	48
4.2.2	Source Code Training Data.....	48
4.2.3	Source Code Preprocessing.....	48
4.2.4	Source Code Classification .....	49
4.2.5	Source Code Validasi.....	50
4.3	Hasil dan Pembahasan Evaluasi Program .....	51
4.3.1	Evaluasi Program Menggunakan Confusion Matrix.....	51
4.3.2	Hasil Pengujian .....	51
<b>BAB V</b>	<b>KESIMPULAN.....</b>	<b>54</b>
5.1	Kesimpulan.....	54
5.2	Saran.....	54
<b>DAFTAR PUSTAKA .....</b>		<b>56</b>

## DAFTAR TABEL

Tabel 2. 1 Confusion Matrix .....	16
Tabel 3. 1 Data Tweet Sebelum Case Folding.....	22
Tabel 3. 2 Data Setelah Case Folding .....	22
Tabel 3. 3 Hasil Tokenizing .....	23
Tabel 3. 4 Hasil Filtering .....	24
Tabel 3. 5 Hasil Stemming .....	26
Tabel 3. 6 Proses Pelabelan Kata .....	28
Tabel 3. 7 Proses Training .....	30
Tabel 3. 8 Dokumen Kategori Negatif.....	32
Tabel 3. 9 Dokumen Kategori Positif .....	34
Tabel 4. 1 Hasil Confusion Matrix.....	52
Tabel 4. 2 Hasil Pengujian Confusion Matrix.....	53

## DAFTAR GAMBAR

Gambar 3. 1 Flowchart Sistem.....	38
Gambar 3. 2 Diagram Alur Preprocessing.....	39
Gambar 3. 3 Proses Klasifikasi.....	40
Gambar 3. 4 Antar Muka Halaman Awal.....	41
Gambar 3. 5 Antar Muka Halaman Hasil.....	41
Gambar 3. 6 Program Download Tweet.....	42
Gambar 4. 1 Data Hasil Pengumpulan Tweet.....	44
Gambar 4. 2 Data Berkategori Positif.....	45
Gambar 4. 3 Data Berkategori Negatif.....	45
Gambar 4. 4 Halaman Utama.....	46
Gambar 4. 5 Halaman Hasil.....	47
Gambar 4. 6 Halaman Validasi.....	47
Gambar 4. 7 Source Code Pengumpulan Data Tweet.....	48
Gambar 4. 8 Source Code Training Data.....	48
Gambar 4. 9 Source Code Preprocessing.....	49
Gambar 4. 10 Source Code Klasifikasi.....	49
Gambar 4. 11 Source Code Validasi.....	51

## INTISARI

Twitter merupakan layanan jejaring sosial dan microblog daring yang memungkinkan penggunaannya untuk mengirim dan membaca pesan berbasis teks hingga 280 karakter. Twitter seringkali digunakan untuk menyampaikan keluhan terhadap suatu produk, tidak luput juga tentang kualitas produk perusahaan layanan telekomunikasi seluler. Melalui cuitan yang diunggah oleh pengguna tersebut dapat dilakukan penggalian informasi mengenai gambaran dari suatu produk tertentu.

Penelitian ini menggunakan algoritma naive bayes untuk melakukan klasifikasi tingkat kepuasan pengguna atas kualitas layanan perusahaan telekomunikasi. Data yang akan digunakan merujuk ke perusahaan yang paling banyak penggunaannya. Data tersebut akan dibagi menjadi kategori positif untuk komentar yang diklasifikasikan positif dan kategori negatif untuk komentar yang diklasifikasikan negatif.

Data yang digunakan untuk proses training adalah 500 data yang terdiri dari 250 komentar positif dan 250 komentar negatif. Penelitian ini dimulai dengan preprocessing data (*case folding, tokenization, filtering, stemming*), proses selanjutnya adalah training data dan proses klasifikasi. Berdasarkan evaluasi menggunakan *confusion matrix* diperoleh hasil akurasi 72%, presisi 70%, dan recall 78%. Dari hasil tersebut, algoritma naive bayes dapat digunakan untuk melakukan klasifikasi komentar twitter terhadap layanan telekomunikasi di Indonesia.

**Kata kunci : Klasifikasi, Algoritma Naive Bayes, Twitter, Confusion Matrix.**

## **ABSTRACT**

*Twitter is an online social networking and microblogging service that allows users to send and read text-based messages of up to 280 characters. Twitter is often used to submit complaints about a product, not to mention the quality of products of cellular telecommunications service companies. Through tweets uploaded by these users, information about the description of a particular product can be extracted.*

*This study uses the Naive Bayes algorithm to classify the level of user satisfaction on the service quality of telecommunications companies. The data to be used refers to the companies with the most users. The data will be divided into positive categories for comments classified as positive and negative categories for comments classified as negative.*

*The data used for the training process is 500 data consisting of 250 positive comments and 250 negative comments. This research begins with preprocessing data (case folding, tokenization, filtering, stemming), the next process is training data and classification process. Based on the evaluation using the Confusion Matrix, the results obtained are 72% accuracy, 70% precision, and 78% recall. From these results, the Naive Bayes algorithm can be used to classify Twitter comments on telecommunications services in Indonesia.*

***Keyword : Classification, Naive Bayes Algorithm, Twitter, Confusion Matrix.***