

**ANALISIS SENTIMEN PADA REVIEW TOKO ONLINE  
MENGUNAKAN METODE NAÏVE BAYES DAN SELEKSI  
FITUR CHI-SQUARE**

**SKRIPSI**



disusun oleh

**Try Satria Amanatullah**

**17.21.1060**

**PROGRAM SARJANA  
PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2019**

**ANALISIS SENTIMEN PADA REVIEW TOKO ONLINE  
MENGUNAKAN METODE NAÏVE BAYES DAN SELEKSI  
FITUR CHI-SQUARE**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Teknik Informatika



disusun oleh

**Try Satria Amanattullah**

**17.21.1060**

**PROGRAM SARJANA  
PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2019**

# PERSETUJUAN

## SKRIPSI

### ANALISIS SENTIMEN PADA REVIEW TOKO ONLINE MENGUNAKAN METODE NAÏVE BAYES DAN SELEKSI FITUR CHI-SQUARE

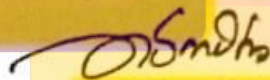
yang dipersiapkan dan disusun oleh

**Try Satria Amanattullah**

**17.21.1060**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 29/04/2019

**Dosen Pembimbing,**



**Windha Mega Pradnya D, M.Kom**

**NIK. 190302185**

**PENGESAHAN**

**SKRIPSI**

**ANALISIS SENTIMEN PADA REVIEW TOKO ONLINE  
MENGUNAKAN METODE NAÏVE BAYES DAN SELEKSI FITUR  
CHI-SQUARE**

yang dipersiapkan dan disusun oleh

**Try Satria Amanattullah**

**17.21.1060**

telah dipertahankan di depan Dewan Penguji  
pada tanggal 26/04/ 2019

**Susunan Dewan Penguji**

**Nama Penguji**

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

**NIK. 190302038**


**Bambang Sudaryatno, Drs., M.M.**

**NIK. 190302029**

**Windha Meza Pradnya D., M.Kom**

**NIK. 190302185**

**Tanda Tangan**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 26/04/2019

**DEKAN FAKULTAS ILMU KOMPUTER**



## 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 tanggung jawab saya pribadi.

Yogyakarta, 5 Mei 2019



Try Satria Amanatullah

NIM. 17.21.1060

## MOTTO

*Working in the silence, let your success be your noise*

(Frank Ocean)

*If you don't understand my silence, you will not understand my words.*

(Unknown)

*Do you want to change yourself? Then remain completely silent inside the silence sea*

(Sri Chinmoy)

*Pain changes people*

(Unknown)

*If you want to be strong learn enjoy being alone*

(Unkown)

Menderita di dalam diam, mendengar di dalam kesunyian, melihat di dalam kegelapan, dan hancurkan musuhmu didalam ketidak pengetahuan.

(Try Satria Amanattullah)

## PERSEMBAHAN

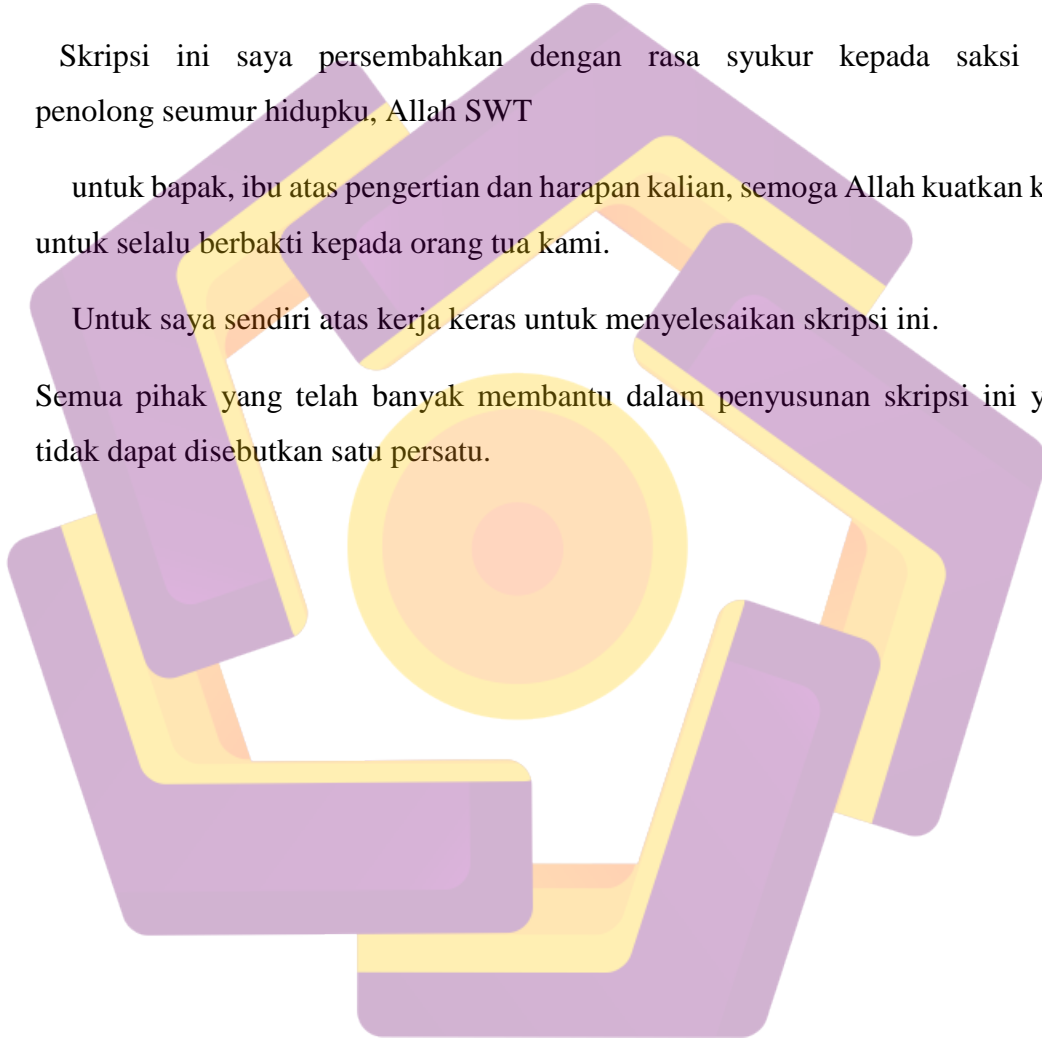
Alhamdulillah segala puji syukur atas kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya sehingga karya ini dapat terselesaikan dengan sebaik – baiknya, tidak lepas dari bantuan dan dukungan berbagai pihak.

Skripsi ini saya persembahkan dengan rasa syukur kepada saksi dan penolong seumur hidupku, Allah SWT

untuk bapak, ibu atas pengertian dan harapan kalian, semoga Allah kuatkan kami untuk selalu berbakti kepada orang tua kami.

Untuk saya sendiri atas kerja keras untuk menyelesaikan skripsi ini.

Semua pihak yang telah banyak membantu dalam penyusunan skripsi ini yang tidak dapat disebutkan satu persatu.



## KATA PENGANTAR

Puji syukur kepada Tuhan Yang Maha Esa karena ridho dan karunianya sehingga penulis dapat menyelesaikan tugas akhir yang berjudul "Analisis Sentimen pada Review Toko Online Menggunakan Metode Naïve Bayes dan Seleksi Fitur Chi-Square" Pada kesempatan ini, penulis ingin menyampaikan ucapan terima kasih kepada pihak-pihak yang telah membantu proses penyelesaian Tugas Akhir ini. Diantaranya adalah :

1. Bapak Prof. Dr. M. Suyanto, MM selaku rector Universitas Amikom Yogyakarta.
2. Bapak Sudarmawan, M.T. Selaku ketua program studi S1 Informatika.
3. Ibu Windha Mega Pradnya D, M.Kom, selaku dosen pembimbing yang telah memandu dan mengarahkan dalam pembuatan skripsi ini.
4. Seluruh dosen, staff maupun karyawan Universitas Amikom Yogyakarta.
5. Kedua orang tua saya dan keluarga yang telah mendoakan dan mendukung selama ini.
6. Teman-teman di Universitas Amikom Yogyakarta yang telah memberikan dukungan.
7. Kepada diri saya sendiri yang telah bekerja keras untuk menyelesaikan skripsi ini

Dalam penulisan ini, masih banyak kekurangan dalam penyusunannya. Untuk itu, penulis mengharapkan kritik dan saran yang membangun dari berbagai pihakpihak lain

Yogyakarta, April 2019

Penulis



## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	i
<b>PERSETUJUAN</b> .....	ii
<b>PENGESAHAN</b> .....	iii
<b>PERNYATAAN</b> .....	iv
<b>MOTTO</b> .....	v
<b>PERSEMBAHAN</b> .....	vi
<b>KATA PENGANTAR</b> .....	vii
<b>DAFTAR ISI</b> .....	viii
<b>DAFTAR TABEL</b> .....	xi
<b>DAFTAR GAMBAR</b> .....	xiii
<b>INTISARI</b> .....	xv
<b>ABSTRACT</b> .....	xvi
<b>BAB I</b> .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Penelitian.....	2
1.3 Batasan Penelitian .....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	4
1.6 Metodologi Penelitian .....	4
1.6.1 Metode Pengumpulan data.....	4
1.6.2 Analisis dan Perancangan .....	4
1.6.3 Metode Implementasi.....	5
1.6.4 Pengujian.....	5
1.7 Sistematika Penulisan.....	5
<b>BAB II</b> .....	7
4.1 Kajian Pustaka.....	7
4.2 Preprocessing .....	8
4.2.1 Tokenization .....	8
4.2.2 Stopword Removal.....	8

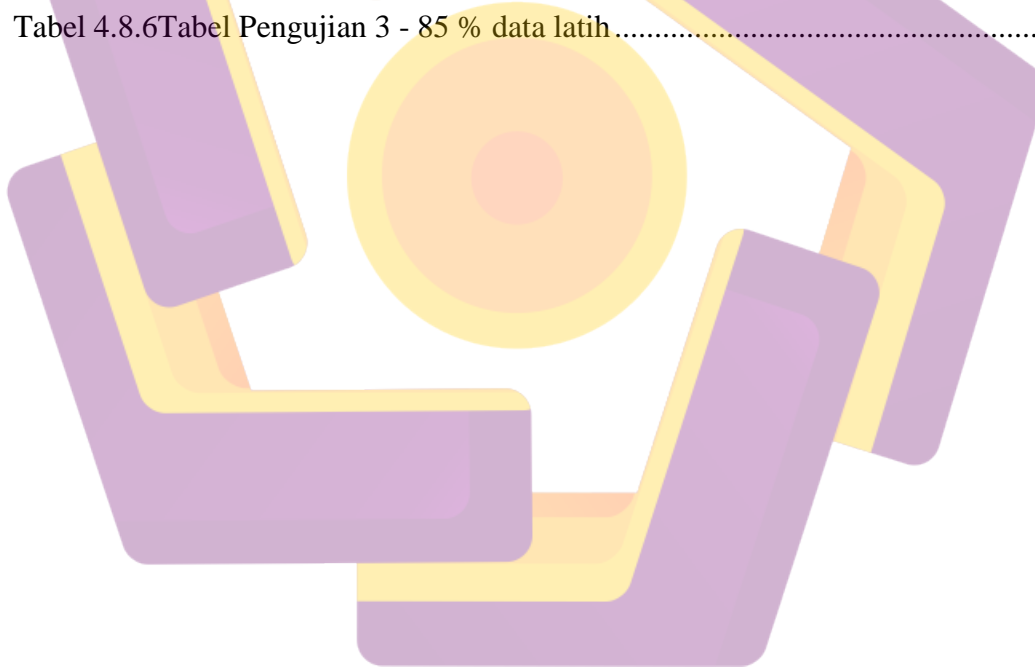
4.2.3	Stemming .....	9
4.3	Chi-Square.....	9
4.4	Naive Bayes Classifier .....	10
4.5	XAMPP .....	10
4.6	Basis Data.....	11
4.6.1	MySQL .....	11
4.7	PHP.....	12
4.8	HTML.....	12
4.9	CSS.....	13
4.9.1	<i>Content Management System (CMS)</i> .....	14
4.10	Flowchart.....	15
4.11	Unified Modeling Language .....	16
4.11.1	Use Case Diagram .....	16
4.11.2	Sequence Diagram .....	17
4.11.3	Class Diagram.....	17
4.12	Confussion Matrix.....	18
<b>BAB III</b>	.....	20
5.2	Gambaran Umum Jalan Proses Penelitian .....	20
5.3	Analisis Kebutuhan Sistem .....	21
5.3.1	Kebutuhan Fungsional .....	21
5.3.2	Kebutuhan Non Fungsional .....	22
5.4	Analisa Algoritma .....	22
5.4.1	Pengumpulan Data .....	22
5.4.2	Pelabelan Manual .....	24
5.4.3	Tokenizing .....	26
5.4.4	Stopword Removal.....	28
5.4.5	Stemming .....	31
5.4.6	Chi-Square .....	34
5.4.7	Klasifikasi Naïve Bayes .....	37
5.5	Perancangan Sistem.....	45
5.5.1	Diagram Use case.....	45

5.5.2	Skenario use case .....	46
5.5.3	Sequence Diagram .....	54
5.5.4	Class Diagram .....	58
5.6	Perancangan Database .....	59
5.6.1	Diagram ERD .....	59
5.6.2	Tabel relasi .....	59
5.6.3	Perancangan Tabel .....	60
5.7	Perancangan Interface .....	61
5.7.1	Interface Ulasan .....	61
5.7.2	Interace Tokenizing .....	62
5.7.3	Interface Stopword Removal .....	62
5.7.4	Interface stemming .....	63
5.7.5	Interface Chi Square .....	63
5.7.6	Interface Naïve Bayes .....	64
5.7.7	Interface Kesimpulan .....	64
<b>BAB IV</b>	.....	65
6.8	Implementasi .....	65
6.8.1	Implementasi Database .....	65
6.8.2	Implementasi Sistem .....	68
6.8.3	Implementasi Interface .....	85
6.9	Evaluasi Hasil .....	94
<b>BAB V</b>	.....	98
7.2	Kesimpulan .....	98
7.3	Saran .....	98
<b>DAFTAR PUSTAKA</b>	.....	99

## DAFTAR TABEL

Tabel 2.10.1 Tabel Flowchart .....	15
Tabel 2.11.1 Tabel Use Case .....	16
Tabel 2.11.2 Tabel Sequence Diagram .....	17
Tabel 2.11.3 Tabel Class Diagram.....	17
Tabel 2.12.1Tabel Confusion Matrix.....	18
Tabel 3.3.1 pengumpulan data .....	22
Tabel 3.3.2 label manual .....	24
Tabel 3.3.3 tokenizing.....	26
Tabel 3.3.4 stopword removal .....	29
Tabel 3.3.5 stemming.....	32
Tabel 3.3.6 chi square .....	34
Tabel 3.3.7 data uji.....	37
Tabel 3.3.8 Hasil perhitungan nilai posterior ulasan 1 .....	39
Tabel 3.3.9 Hasil perhitungan nilai posterior ulasan 2 .....	40
Tabel 3.3.10 Hasil perhitungan nilai posterior ulasan 3 .....	42
Tabel 3.3.11 Hasil perhitungan nilai posterior ulasan 4 .....	43
Tabel 3.3.12 Hasil klasifikasi.....	44
Tabel 3.3.13 confusion Matrix.....	45
Tabel 3.4.1 Skenario use case input ulasan.....	46
Tabel 3.4.2 Skenario use case label manual.....	47
Tabel 3.4.3 Skenario use case tokenizing .....	48
Tabel 3.4.4 Skenario use case stopword removal .....	48
Tabel 3.4.5 Skenario use case stemming .....	49
Tabel 3.4.6 Skenario use case chi square.....	50
Tabel 3.4.7 Skenario use case naive bayes .....	51
Tabel 3.4.8 Skenario use case hasil kesimpulan .....	52
Tabel 3.4.9 Skenario use case akurasi.....	53

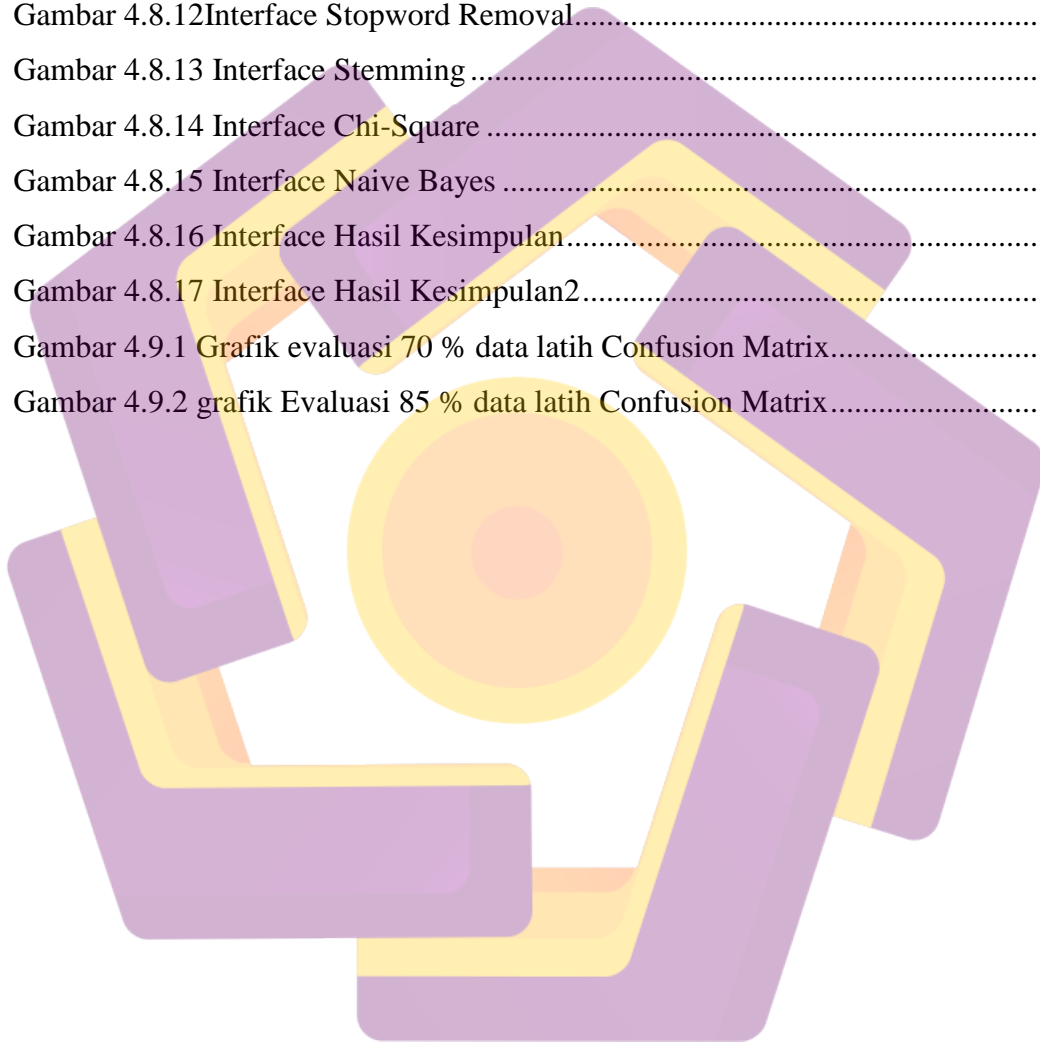
Tabel 3.5.1 perancangan tabel ulasan .....	60
Tabel 3.5.2 perancangan tabel tokenizing.....	60
Tabel 3.5.3 perancangan tabel stopwords.....	60
Tabel 3.5.4 perancangan tabel stemming.....	60
Tabel 3.5.5 perancangan tabel fitur chi square .....	61
Tabel 3.5.6 perancangan tabel hasil klasifikasi.....	61
Tabel 4.8.1Tabel Pengujian 1 - 70 % data latih.....	94
Tabel 4.8.2 Tabel Pengujian 2 - 70 % data latih.....	94
Tabel 4.8.3 Tabel Pengujian 3 - 70 % data latih.....	95
Tabel 4.8.4Tabel Pengujian 1 - 85 % data latih.....	96
Tabel 4.8.5Tabel Pengujian 2 - 85 % data latih.....	96
Tabel 4.8.6Tabel Pengujian 3 - 85 % data latih.....	96



## DAFTAR GAMBAR

Gambar 3.2.1 Diagram alir Proses Penelitian.....	20
Gambar 3.5.1 Use Case.....	46
Gambar 3.5.2 Sequence input ulasan.....	54
Gambar 3.5.3 Sequence label manual.....	55
Gambar 3.5.4 Sequence tokenizing.....	55
Gambar 3.5.5 Sequence stopword removal.....	56
Gambar 3.5.6 Sequence stemming.....	56
Gambar 3.5.7 Sequence chi square.....	57
Gambar 3.5.8 Sequence naive bayes.....	57
Gambar 3.5.9 Sequence akurasi.....	58
Gambar 3.5.10 Class Diagram.....	58
Gambar 3.6.1 Diagram ERD.....	59
Gambar 3.6.2 Relasi Tabel.....	59
Gambar 3.7.1 Rancangan interface ulasan.....	61
Gambar 3.7.2 rancangan interface tokenizing.....	62
Gambar 3.7.3 rancangan interface stopword.....	62
Gambar 3.7.4 rancangan interface stemming.....	63
Gambar 3.7.5 rancangan interface chi square.....	63
Gambar 3.7.6 rancangan interface naive bayes.....	64
Gambar 3.7.7 rancangan interface kesimpulan.....	64
Gambar 4.8.1 Database.....	65
Gambar 4.8.2 Tabel dokumen.....	66
Gambar 4.8.3 Tabel Tokenizer.....	66
Gambar 4.8.4 Tabel Stopword.....	66
Gambar 4.8.5 Tabel Stemming.....	67
Gambar 4.8.6 Tabel Fitur.....	67
Gambar 4.8.7 Tabel hasil Klasifikasi.....	68

Gambar 4.8.8 Interface Ulasan .....	85
Gambar 4.8.9 Interface Data Latih.....	86
Gambar 4.8.10 Interface Data Uji.....	87
Gambar 4.8.11 Interface Tokenizing .....	88
Gambar 4.8.12Interface Stopword Removal.....	89
Gambar 4.8.13 Interface Stemming .....	90
Gambar 4.8.14 Interface Chi-Square .....	91
Gambar 4.8.15 Interface Naive Bayes .....	92
Gambar 4.8.16 Interface Hasil Kesimpulan.....	93
Gambar 4.8.17 Interface Hasil Kesimpulan2.....	93
Gambar 4.9.1 Grafik evaluasi 70 % data latih Confusion Matrix.....	95
Gambar 4.9.2 grafik Evaluasi 85 % data latih Confusion Matrix.....	97



## INTISARI

Analisis sentiment biasanya digunakan untuk menggali informasi dari kumpulan data berupa pengetahuan yang selama ini tidak diketahui secara manual. Salah satu metode yang bisa digunakan untuk menyelesaikan masalah adalah menggunakan metode *Naïve Bayes* dan seleksi fitur *Chi-Square*.

Pada saat ini penjualan mulai berkembang ke era digital dimana semuanya bisa dijual secara online, dengan itu dibutuhkan aplikasi yang dapat menganalisis masing-masing review masyarakat secara objektif.

Penelitian ini Menggunakan data yang diambil dari data review pada toko online yang kemudian dianalisis menggunakan metode *naïve bayes* dan seleksi fitur *chi-square* yang akan menghasilkan hasil kesimpulan berupa frekuensi Fitur dari yang tertinggi ke yang terendah

**Kata Kunci** : Analisis Sentimen, *Naïve Bayes*, *Chi-Square*



## ***ABSTRACT***

*Sentiment analysis is usually used to extract information from the form of the data set of knowledge that had been unknown manually. One method that can be used to solve the problem is using Naïve Bayes method and chi-square feature selection.*

*At the moment sales are starting to develop into the digital era where everything can be sold online, it need an applications that can assess each Review objectively.*

*This study uses data taken from review data on online stores which is then analyzed using the naïve Bayes method and chi-square feature selection that will produce the results of conclusions in the form of frequency from the highest to the lowest.*

***Keywords: Sentiment analysis, Naïve Bayes, Chi-Square***

