

**ANALISIS SENTIMEN PADA REVIEW PERGURUAN TINGGI DENGAN
TEKS BAHASA INDONESIA MENGGUNAKAN ALGORITMA NAIVE
BAYES**

SKRIPSI



disusun oleh

Ilham Wiyono

16.11.0497

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

**ANALISIS SENTIMEN PADA REVIEW PERGURUAN TINGGI DENGAN
TEKS BAHASA INDONESIA MENGGUNAKAN ALGORITMA NAIVE
BAYES**

SKRIPSI

untuk memenuhi sebagaimana persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



disusun oleh

Ilham wiyono

16.11.0497

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

PERSETUJUAN

SKRIPSI

**ANALISIS SENTIMEN PADA REVIEW PERGURUAN TINGGI DENGAN
TEKS BAHASA INDONESIA MENGGUNAKAN ALGORITMA NAIVE
BAYES**

yang dipersiapkan dan disusun oleh

Ilham Wiyono

16.11.0497

telah ditetujui oleh Dosen Pembimbing Skripsi
Pada tanggal **Oktober 2021**

Dosen Pembimbing,

Ike Verawati, M.Kom

NIK. 190302237

PENGESAHAN

SKRIPSI

**ANALISIS SENTIMEN PADA REVIEW PERGURUAN TINGGI DENGAN
TEKS BAHASA INDONESIA MENGGUNAKAN ALGORITMA NAIVE
BAYES**

yang dipersiapkan dan disusun oleh

Ilham Wiyono

16.11.0497

Telah dipertimbangkan di depan Dewan Penguji
pada tanggal **Oktober 2021**

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

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

**Sumarni Adi, S.Kom, M.Cs NIK.
190302256**

**Anna Baita M.Kom NIK.
190302290**

Skripsi ini telah diterima sebagai salah satu persyaratan
Untuk memperoleh gelar Sarjana Komputer
Tanggal **Oktober 2021**

DEKAN FAKULTAS ILMU KOMPUTER

**Hanif Al Fatta, S.Kom, M.Kom
NIK. 190302096**

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, November 2020

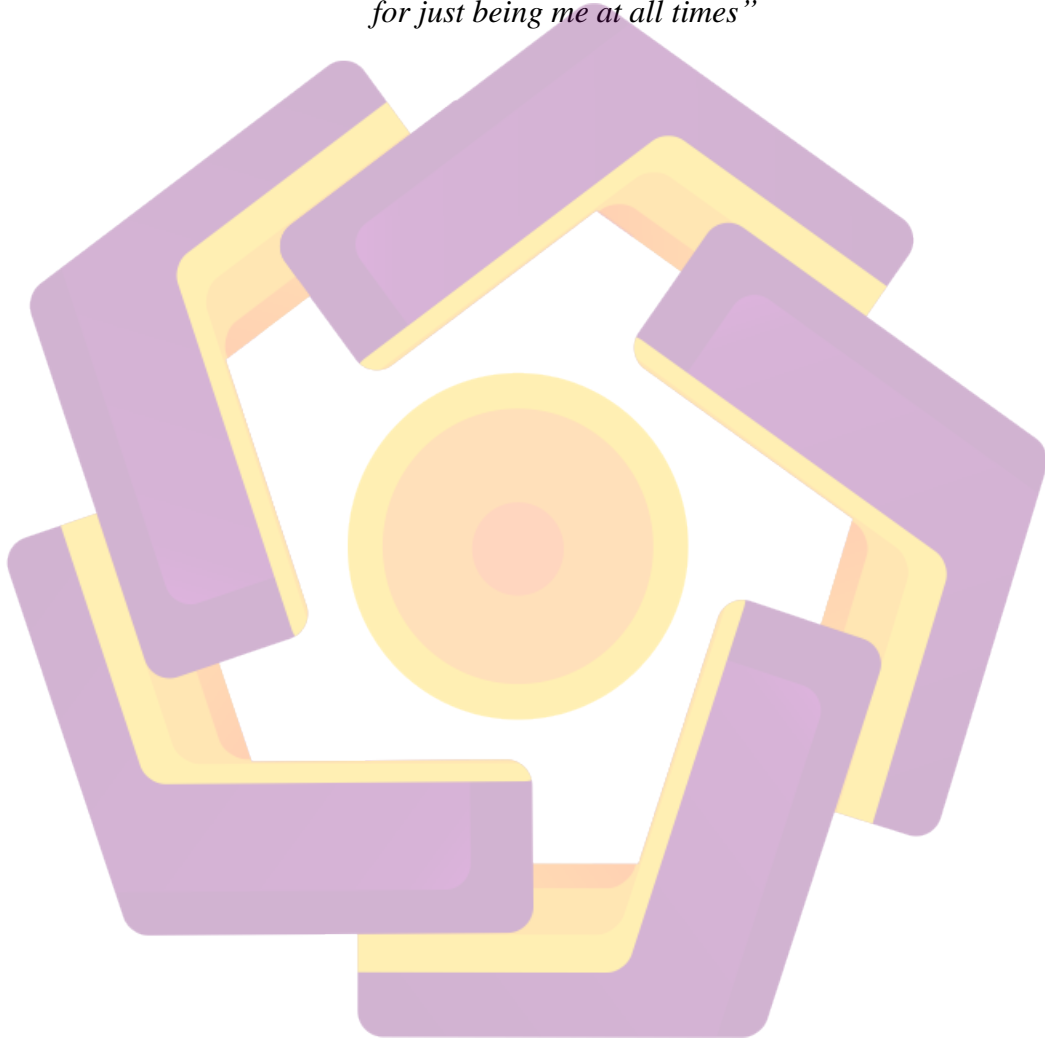


Ilham wiyono

16.11.0496

MOTTO

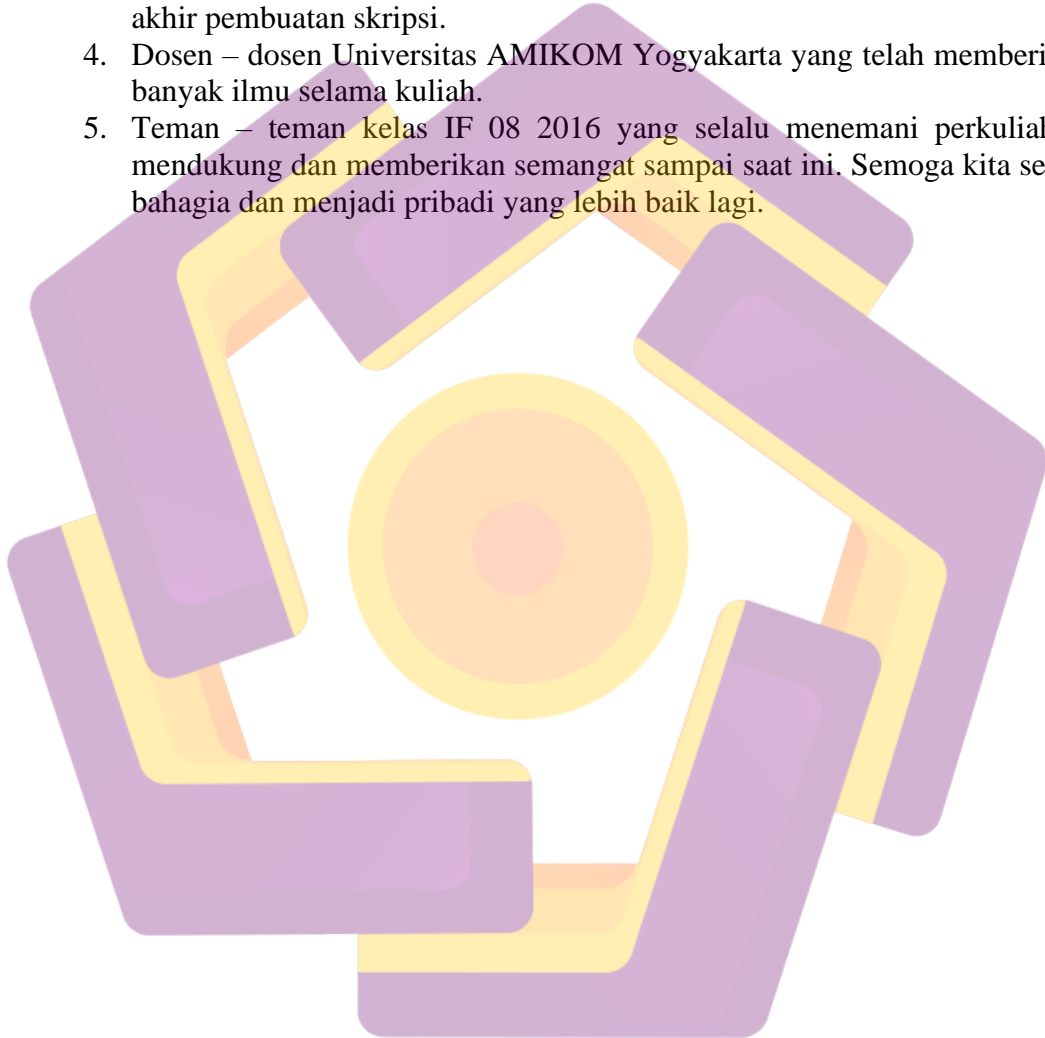
*“Last but not least, I wanna thank me
for believing in me, for doing all this hard work
for having no days off, for never quitting
for always being a giver
And tryna give more than I recieve
I wanna thank me for tryna do more right than wrong
for just being me at all times”*



PERSEMBAHAN

Saya mempersembahkan skripsi ini kepada semua pihak yang terlibat secara langsung maupun tidak langsung dalam proses pembuatan skripsi.

1. Tuhan Yang Maha Esa yang memberikan segala nikmat dan kasih sayangnya sampai sejauh ini.
2. Kedua orang tua saya dan keluarga, yang selalu mendoakan, selalu menyemangati dan memberikan uang jajan kepada saya.
3. Ibu Ike Verawati, M.Kom yang telah membimbing saya dari awal sampai akhir pembuatan skripsi.
4. Dosen – dosen Universitas AMIKOM Yogyakarta yang telah memberikan banyak ilmu selama kuliah.
5. Teman – teman kelas IF 08 2016 yang selalu menemani perkuliahan, mendukung dan memberikan semangat sampai saat ini. Semoga kita selalu bahagia dan menjadi pribadi yang lebih baik lagi.



KATA PENGANTAR

Puji dan syukur panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan rahmat, hidayah dan kekuatan sehingga saya dapat menyelesaikan skripsi yang berjudul Analisis Sentimen Terhadap Tingkat Kepuasan Pengguna Jasa Antar Brang JNE Pada Twitter Menggunakan Algoritma Support Vector Machine (SVM).

Skripsi ini saya buat guna menyelesaikan studi jenjang Starta Satu (S1) pada program studi Informatika fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta. Selain itu juga merupakan suatu bukti bahwa mahasiswa telah menyelesaikan kuliah jenjang program strata satu dan untuk memperoleh gelar Sarjana Komputer. Dengan selesainya skripsi ini, maka pada kesempatan ini saya mengucapkan terima kasih 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. Ibu Ike Verawati, M.Kom selaku dosen pembimbing yang selalu bijaksana dalam memberikan bimbingan, nasehat serta waktunya selama pembuatan skripsi ini.
4. Dosen penguji Ibu Ike Verawati M.Kom Ibu Sumarni Adi, S.Kom, M.Cs, Ibu Anna Baita M.Kom, dan segenap Dosen serta Karyawan Universitas AMIKOM Yogyakarta yang telah berbagi ilmu dan pengalamannya.
5. Kedua orang tua saya yang telah mendoakan, mendukung dan memberikan semangat kepada saya.
6. Teman – teman IF 08 2016 yang telah menemani selama proses perkuliahan.
7. Semua pihak yang telah membantu baik secara langsung maupun tidak langsung.
8. Serta orang yang telah meninggalkan ku.

Semoga Allah SWT memberikan balasan yang lebih kepada semua yang telah ikut membantu saya hingga menyelesaikan skripsi ini. Demi perbaikan selanjutnya, saran dan kritik membangun diterima dengan senang hati dan terima kasih. Semoga skripsi ini dapat bermanfaat bagi saya dan kita semua.

Yogyakarta, 20 Oktober 2021

Ilham WIyono

DAFTAR ISI

<u>PERSETUJUAN</u>	Error! Bookmark not defined.i.
<u>PENGESAHAN</u>	Error! Bookmark not defined.ii.
<u>PERNYATAAN</u>	Error! Bookmark not defined.iii.
<u>MOTTO</u>	Error! Bookmark not defined.v.
<u>PERSEMBAHAN</u>	Error! Bookmark not defined.
<u>KATA PENGANTAR</u>	Error! Bookmark not defined.
<u>DAFTAR ISI</u>	viiError! Bookmark not defined.
<u>DAFTAR GAMBAR</u>	Error! Bookmark not defined.
<u>DAFTAR TABEL</u>	xiError! Bookmark not defined.
<u>INTISARI</u>	Error! Bookmark not defined.
<u>ABSTRACT</u>	Error! Bookmark not defined.
BAB I	1
1 PENDAHULUAN	1
2 Latar Belakang	1
3 Rumusan Masalah	2
4 Batasan Masalah	2
5 Tujuan penulisan	3
6 Manfaat Penulisan	4
7 Metode Penelitian	4
8 Metode Pengumpulan data	4
12 Metode analisis	5
13 Metode Implementasi	6
14 Metode testing	7
15 Sistematika Penulisan	7
BAB II	9
16 LANDASAN TEORI	9
16.1 Tinjauan Pustaka	9
17 Pengertian Data Mining	14
18 Pengertian Sentiment Analysis	15

19	Pengertian Text Mining	16
20	Aplikasi Text Mining.....	18
21	Manfaat Text Mining	19
22	Tantang tantangan dan masalah Text Mining	20
23	Pengertian Pre-Processing.....	20
24	Naïve bayes.....	21
25	Confusion Matrix	24
26	Flowchart	26
27	System Flowchart (Diagram Alir Sistem)	26
28	Program Flowchart (Diagram Alir Program).....	27
29	Bahasa Pemograman Python	28
BAB III.....		30
METODOLOGI PENELITIAN		30
3.1	Alur penelitian.....	30
3.1.1	Pengumpulan data	30
3.1.2	Preprocessing data	31
1.	Cleaning data.....	31
2.	Case folding	31
3.	Tokenization	31
4.	Stimming	31
3.1.3	Import data.....	32
3.1.4	Pelabelan data	32
3.1.5	Pengujian data.....	32
3.2	Alat penelitian	32
3.2.1	hadrware.....	32
3.2.2	Software	33
3.3	Bahan Penelitian	34
3.3.1	Data tweet.....	34
3.3.2	Bahasa Pemrograman python.....	34
3.3.3	Bahasa pemrograman php	35
BAB IV		36
IMPLEMENTASI DAN PEMBAHASAN		36
4.1	Implementasi Pengumpulan Data Tweet	36
4.2	Diagram Alur Preprocessing	38

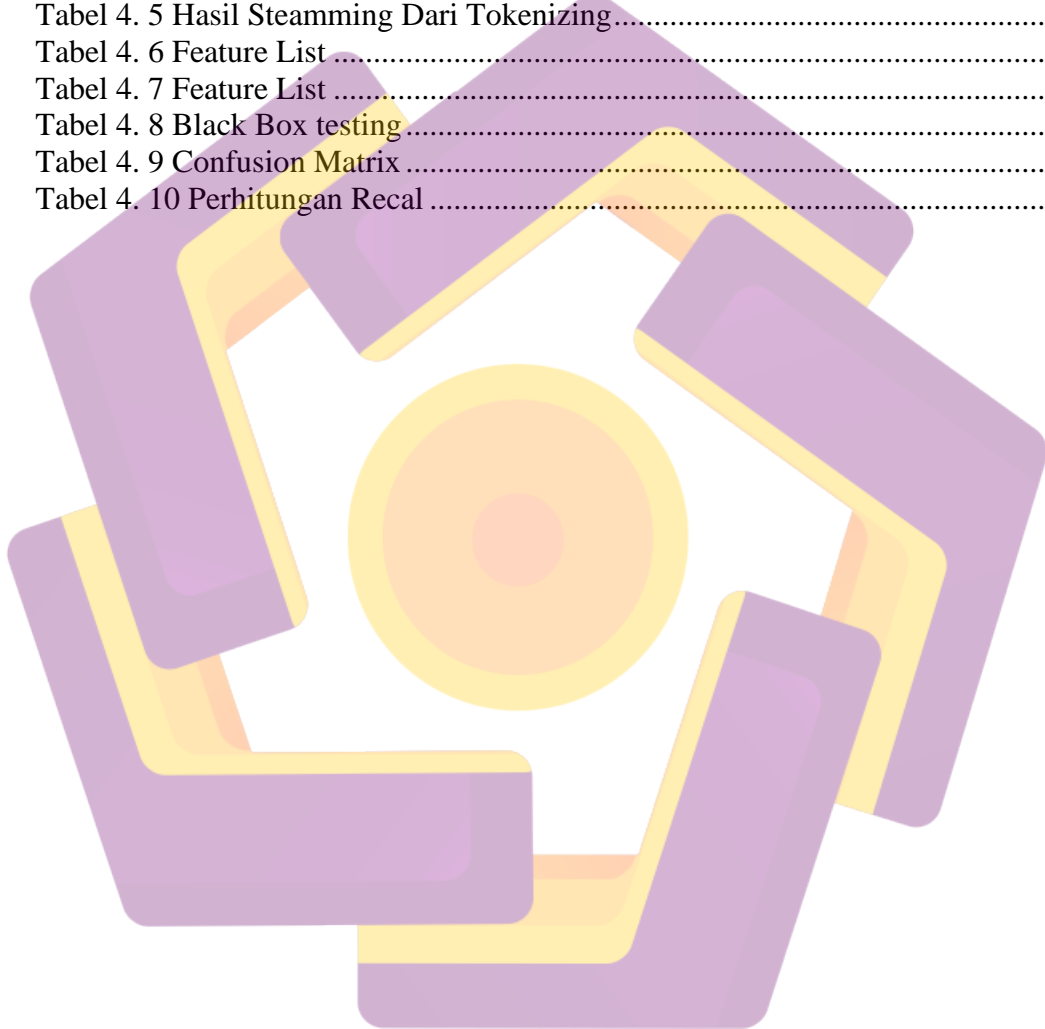
4.3	Cleaning Data	39
4.4	Case Folding.....	41
4.5	Tokenization.....	41
4.6	Steaming.....	42
4.7	Klasifikasi Naïve bayes Clasifier	44
4.8	Pelebelan Data Training	48
4.9	Proses olah data universitas	48
4.4	Proses Import data	51
4.5	Proses analisis Sentimen	53
4.6	Proses pengujian.....	54
4.6.1	Black Box testing.....	59
4.7	Pembuatan Interface	60
4.7.1	Halaman login	60
4.7.2	Halaman Utama	61
4.7.3	Halaman Data pengguna	61
4.7.5	Halaman Data Universitas	62
4.7.6	Halaman Data Tweet	62
4.7.7	Halaman hasil sentiment	63
4.7.8	Halaman pengujian akurasi.....	64
4.7.9	Halaman Hasil Klasifikasi.....	66
BAB V	70
5.1	Kesimpulan.....	70
5.2	Saran	70

DAFTAR GAMBAR

Gambar 2. 1 System Flowchart (Diagram Alir System).....	26
Gambar 2. 2 Program Flowchart (Diagram Alir Program).....	27
Gambar 3. 1 Alur Penelitian	30
Gambar 4. 1 Script Crawling Data Twitter	36
Gambar 4. 2 Hasil Crawling Data.....	38
Gambar 4. 3 Diagram Alur Pre-processing.....	38
Gambar 4. 4 Proses Klasifikasi Naïve bayes Clasifier.....	44
Gambar 4. 5 Pelabelan Data.....	48
Gambar 4. 6 Halaman Login.....	61
Gambar 4. 7 Halaman Utama.....	61
Gambar 4. 8 Halaman data pengguna	62
Gambar 4. 9 Halaman data universitas	62
Gambar 4. 10 Halaman data Tweet.....	63
Gambar 4. 11 Halaman hasil sentiment	63
Gambar 4. 12 hasil jumlah data sentiment.....	64
Gambar 4. 13 Grafik hasil analisis sentiment	64
Gambar 4. 14 Halaman hasil sentiment	65
Gambar 4. 15 Pengujian Rasio data	65
Gambar 4. 16 Hasil pengujian.....	69

DAFTAR TABEL

Tabel 2. 1 Perbandingan penelitian terkait dengan sistem yang dibangun	11
Tabel 2. 2 Confusion Matrix	24
Tabel 3. 1 Data tweet	34
Tabel 4. 1 Data Tweet Sebelum Dilakukan Cleaning Data	39
Tabel 4. 2 Data yang telah dilakukan cleaning	40
Tabel 4. 3 Data Tweet Setelah Case Folding	41
Tabel 4. 4 Hasil Tokenization Dari Case Folding.....	42
Tabel 4. 5 Hasil Stemming Dari Tokenizing.....	42
Tabel 4. 6 Feature List	43
Tabel 4. 7 Feature List	47
Tabel 4. 8 Black Box testing	60
Tabel 4. 9 Confusion Matrix	66
Tabel 4. 10 Perhitungan Recal	67



INTISARI

Analisis sentimen merupakan bidang yang saat ini sangat diminati oleh berbagai kalangan. Analisis sentimen dapat dilakukan dengan menggunakan dokumen maupun opini dari media sosial. Salah satu media sosial yang biasanya digunakan sebagai sarana opini adalah media sosial Twitter. Sebelum sebuah teks diklasifikasikan, Perlu melakukan pelabelan kata dimana tujuannya adalah untuk menentukan kata yang termasuk opini dan non opini.

Untuk pelabelan kata dapat menggunakan Hidden Markov Model maupun Rule Based. Adapun metode yang biasa digunakan dalam analisis sentimen adalah metode Naïve Bayes Classifier. Metode ini mengklasifikasikan probabilitas dengan sederhana.

Naïve Bayes Classifier dapat digunakan untuk mengklasifikasikan opini ke dalam opini positif maupun negatif. Selain itu metode ini menggunakan data latih dalam proses klasifikasinya. Klasifikasi yang dihasilkan dari metode Naïve Bayes Classifier tergolong cukup baik. Untuk menguji tingkat keakuratan sistem dalam mengklasifikasikan opini, maka dilakukan pengujian hasil klasifikasi. Dari hasil pengujian diperoleh rata-rata akurasi sebesar 87,1 %. Semakin banyak data latih yang mirip dengan data testing maka hasil klasifikasi akan semakin bagus. Dari hasil tersebut metode Naïve Bayes Classifier dapat digunakan untuk klasifikasi sentimen terhadap data twitter perguruan tinggi.

Kata Kunci: Analisis Sentimen, Naïve Bayes Classifier, akurasi, twitter

ABSTRACT

Sentiment analysis is a field that is currently in great demand by various groups. Sentiment analysis can be done using documents or opinions from social media. One of the social media that is usually used as a means of opinion is Twitter social media. Before a text is classified, it is necessary to do word labeling where the aim is to determine which words are opinion and non opinion.

For word labeling, you can use the Hidden Markov Model or Rule Based. The method commonly used in sentiment analysis is the Naïve Bayes Classifier method. This method classifies the probabilities in a simple way.

Naïve Bayes Classifier can be used to classify opinions into positive or negative opinions. In addition, this method uses training data in the classification process. The classification resulting from the Naïve Bayes Classifier method is quite good. To test the level of accuracy of the system in classifying opinions, the classification results are tested. From the test results obtained an average accuracy of 87.1%. The more training data that is similar to the testing data, the better the classification results will be. From these results, the Naïve Bayes Classifier method can be used to classify sentiments on university twitter data.

Key words: *Sentiment analysis, Naïve Bayes Classifier, accuracy, twitter*

