

**ANALISIS OPINI TERHADAP DC UNIVERSE PADA  
MEDIA SOSIAL TWITTER MENGGUNAKAN  
METODE NAÏVE BAYES**

**SKRIPSI**



disusun oleh

**Andrew Patrick de Fretes**

**16.11.0278**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2020**

**ANALISIS OPINI TERHADAP DC UNIVERSE PADA  
MEDIA SOSIAL TWITTER MENGGUNAKAN  
METODE NAÏVE BAYES**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Informatika



disusun oleh

**Andrew Patrick de Fretes**

**16.11.0278**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2020**

# **PERSETUJUAN**

## **SKRIPSI**

### **ANALISIS OPINI TERHADAP DC UNIVERSE PADA MEDIA SOSIAL TWITTER MENGGUNAKAN METODE NAÏVE BAYES**

yang dipersiapkan dan disusun oleh

**Andrew Patrick de Fretes**

**16.11.0278**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 22 April 2019

**Dosen Pembimbing,**

**Anggit Dwi Hartanto, M.Kom**  
**NIK. 190302163**

# PENGESAHAN

## SKRIPSI

### ANALISIS OPINI TERHADAP DC UNIVERSE PADA MEDIA SOSIAL TWITTER MENGGUNAKAN METODE NAÏVE BAYES

yang dipersiapkan dan disusun oleh

**Andrew Patrick de Fretes**

**16.11.0278**

telah dipertahankan di depan Dewan Penguji  
pada tanggal 19 Maret 2020

**Susunan Dewan Penguji**

**Nama Penguji**

**Anggit Dwi Hartanto, M.Kom**  
**NIK. 190302163**

**Heri Sismoro, S.Kom., M.Kom**  
**NIK. 190302057**

**Sri Ngudi Wahyuni, S.T., M.Kom**  
**NIK. 190302060**

**Tanda Tangan**

---

---

---

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 20 April 2020

**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, 20 April 2020



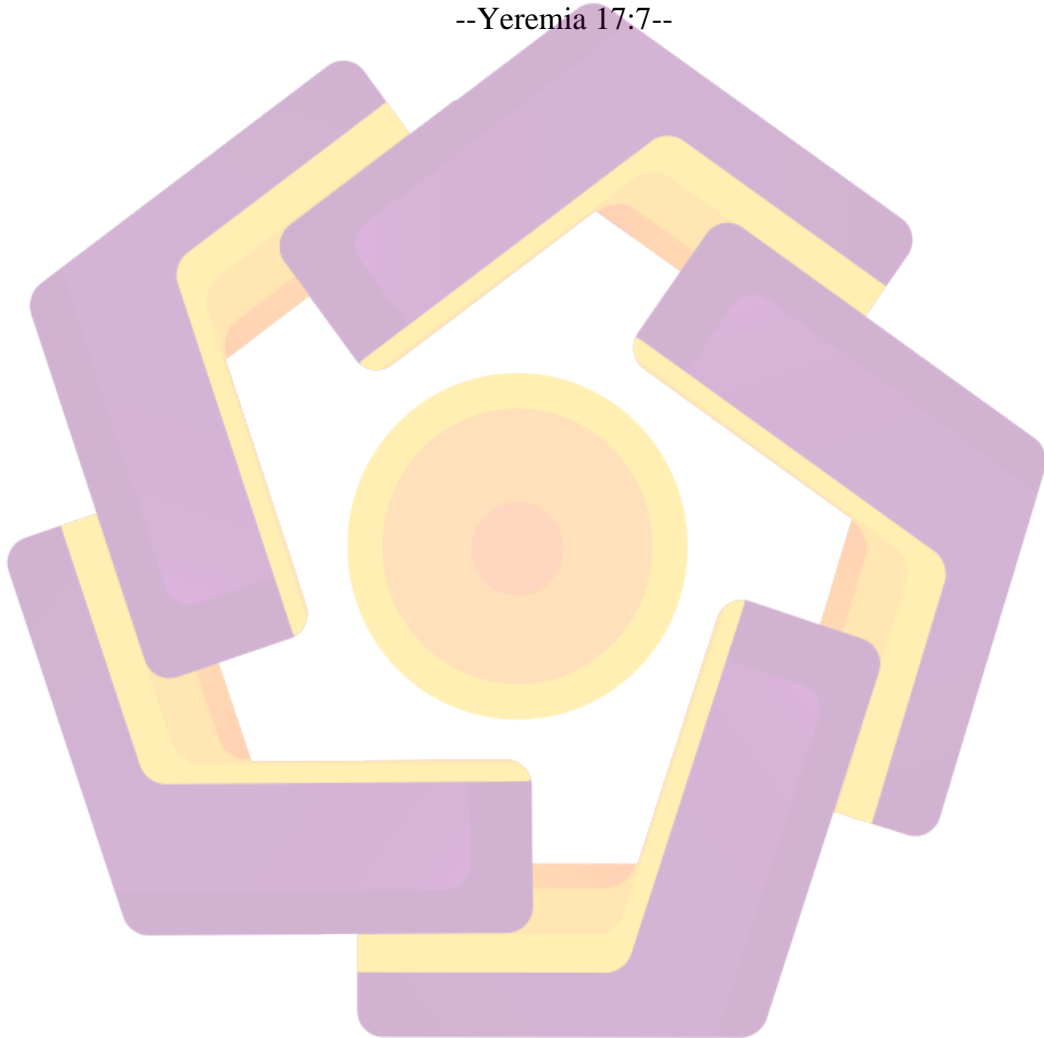
Andrew Patrick de Fretes

NIM. 16.11.0278

## MOTTO

“Diberkatilah orang yang mengandalkan TUHAN, yang menaruh harapannya  
pada TUHAN!”

--Yeremia 17:7--

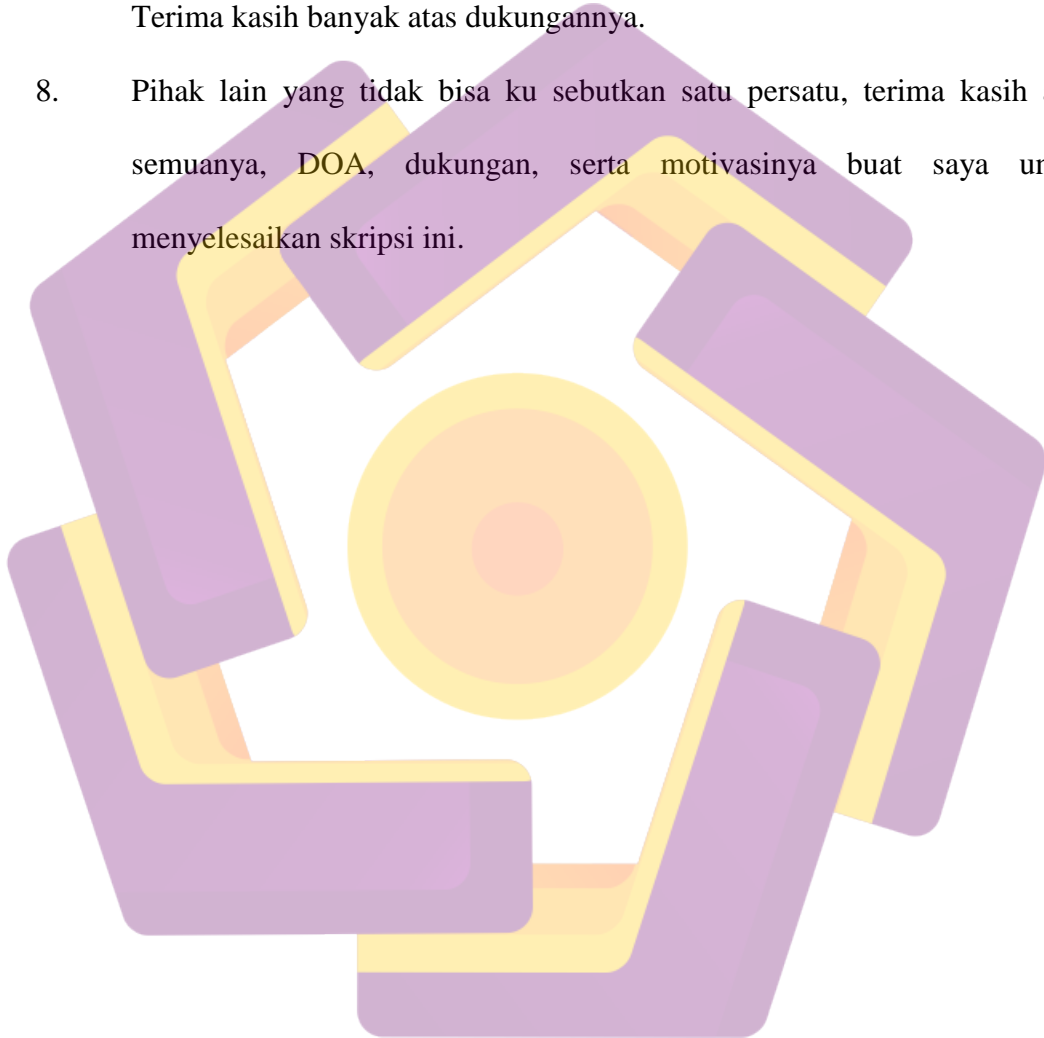


## PERSEMBAHAN

Segala puji syukur saya panjatkan kepada Tuhan Yesus, yang telah memberikan kasih, kesehatan, dan kemudahan kepada saya sehingga dapat menyelesaikan Skripsi ini. Pada halaman persembahan ini, saya juga ingin berterimakasih yang sebesar-besarnya kepada:

1. Bapak dan Mama yang selalu menjaga saya, memberi dukungan, DOA, semangat, dan kepercayaan. Terimakasih sebesar-besarnya atas apa yang telah kalian berikan dan lakukan untuk saya, Tuhan Yesus selalu memberi kebahagiaan dan kesehatan selalu.
2. Keluarga besarku. Ma cici, Ma enda, Ma ros, Bapa koko, kakak herlyn, kakak Jenni, untuk semua yang ada di jayapura dan di ambon yang tidak bisa saya sebutkan satu persatu. Terima kasih atas motivasi, doa, dan kepercayaan untuk saya.
3. Bapak Anggit Dwi Hartanto, M.Kom selaku dosen pembimbing yang telah membimbing dengan sangat sabar dan selalu meluangkan waktunya untuk saya, terimakasih telah memberikan saya kemudahan, saya belajar banyak dari bapak.
4. Sahabat serumahku selama di jogja dari jayapura. Fahmi Ilmawan, Tri Sulestiyawan, Raga Maulana Wahyu Saputra, M. Fitriyanto Saputra. Terimakasih atas semangat dan bantuannya.
5. Sahabatku PT.SANSxVVIVU. Ray reysandi, Denilson Mofu, Exa Adji Putra, Helmi Alfaiz, Zakaria Mursid, R. Gilang Adhitya Pramana, Andika Putra Pratama, Delwyn Talubun. Terimakasih banyak atas semangat dan bantuannya.

6. Teman-teman dari kelas 16 IF 05 yang telah memberikan warna dan cerita dalam perjalanan kuliah saya, Terima kasih banyak.
7. Sahabat-sahabat saya di jayapura. Ajun, Chipin, Gaby, Nuna, Rizky, Hendra, Arif, Diky, Bhimo, Alam, Rani, Uli, Akbar, Bagazt, Chayrul, Fikar. Terima kasih banyak atas dukungannya.
8. Pihak lain yang tidak bisa ku sebutkan satu persatu, terima kasih atas semuanya, DOA, dukungan, serta motivasinya buat saya untuk menyelesaikan skripsi ini.





## KATA PENGANTAR

Puji syukur kepada Tuhan atas segala berkat dan rahmat-Nya sehingga penulis dapat menyelesaikan penulisan skripsi tepat pada waktunya, dengan judul “ANALISIS OPINI TERHADAP DC UNIVERSE PADA MEDIA SOSIAL TWITTER MENGGUNAKAN MEOTDE NAÏVE BAYES”. Skripsi ini merupakan karya ilmiah yang disusun oleh penulis sebagai salah satu syarat kelulusan jenjang pendidikan S1 Informatika di Universitas Amikom Yogyakarta.

Penulis menyadari bahwa dalam penyusunan Skripsi ini terdapat kekurangan, baik dalam penulisan maupun isi, dikarenakan keterbatasan waktu dalam penyusunan Skripsi ini. Dengan rendah hati, penulis mohon maaf dan sangat mengharapkan adanya saran dan kritik yang membangun dari pembaca serta semua pihak yang terkait dalam usaha penyempurnaan skripsi ini.

Pada kesempatan ini penulis ingin menyampaikan rasa hormat dan terima kasih kepada :

1. Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta.
2. Ibu Krisnawati, S.Si, M. T. selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Anggit Dwi Hartanto, M.Kom selaku Dosen Pembimbing yang telah banyak membantu dan membimbing dalam menyelesaikan skripsi ini. Semoga Allah senantiasa melimpahkan rahmat-Nya kepada beliau.
4. Seluruh Dosen Penguji pendadaran yang memberikan saran dan kritik untuk menyempurnakan skripsi ini.
5. Para Dosen di Universitas Amikom Yogyakarta yang telah mendidik dan memberi bekal ilmu yang menjadi modal dasar bagi penulis dalam penulisan skripsi ini.
6. Orang tua penulis yang selalu memberikan doa dan dukungan.
7. Seluruh pihak yang telah membantu dalam kelancaran penulisan skripsi ini baik langsung maupun tidak langsung yang tidak dapat penulis sebutkan satu-persatu.

## DAFTAR ISI

JUDUL .....	I
PERSETUJUAN .....	II
PENGESAHAN .....	III
PERNYATAAN.....	IV
MOTTO .....	V
PERSEMBAHAN.....	VI
KATA PENGANTAR .....	VIII
DAFTAR ISI.....	IX
DAFTAR TABEL.....	XIII
DAFTAR GAMBAR .....	XV
INTISARI.....	XVII
ABSTRACT.....	XIX
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.....	4
1.6 METODE PENELITIAN .....	4
1.6.1 Metode Pengumpulan Data .....	4
1.6.2 Metode Analisis Data.....	5
1.6.3 Metode Perancangan .....	5
1.7 SISTEMATIKA PENULISAN .....	5
BAB II LANDASAN TEORI.....	7

2.1	KAJIAN PUSTAKA.....	7
2.1.1	<i>Penelitian Terdahulu</i> .....	8
2.2	DATA MINING .....	9
2.3	TEXT MINING .....	9
2.4	ANALISIS SENTIMEN .....	10
2.5	TF-IDF .....	10
2.6	PREPROCESSING.....	12
2.6.1	<i>Case Folding</i> .....	12
2.6.2	<i>Cleansing</i> .....	12
2.6.3	<i>Convert Emoji</i> .....	13
2.6.4	<i>Convert Negation</i> .....	13
2.6.5	<i>Tokenizing</i> .....	13
2.6.6	<i>Normalization</i> .....	13
2.6.7	<i>Stopword Removal</i> .....	15
2.6.8	<i>Stemming</i> .....	15
2.6.8.1	<i>Sastrawi Stemmer</i> .....	15
2.7	NAÏVE BAYES CLASSIFIER.....	16
2.8	CONFUSION MATRIX.....	18
2.9	PHP: HYPERTEXT PREPROCESSOR .....	19
2.10	BASIS DATA .....	20
2.11	MYSQL.....	21
BAB III METODE PENELITIAN.....		22
3.1	GAMBARAN UMUM .....	22
3.2	ALAT & BAHAN PENELITIAN .....	23
3.2.1	<i>Alat Penelitian</i> .....	23
3.2.2	<i>Bahan Penelitian</i> .....	24
3.3	ALUR PENELITIAN.....	24
3.4	PERANCANGAN SISTEM.....	25
3.4.1	<i>Perancangan Antar Muka</i> .....	25
a)	<i>Antar Muka Halama Utama</i> .....	25

b)	Antar Muka <i>Import</i> Data latih.....	26
c)	Antar Muka <i>Preprocessing</i> Data latih .....	26
d)	Antar Muka Proses Bobot TF-IDF.....	27
e)	Antar Muka Proses Bobot <i>Naïve Bayes</i> .....	27
f)	Antar Muka <i>Crawling</i> Data uji .....	28
g)	Antar Muka <i>Preprocessing</i> Data uji .....	28
h)	Antar Muka Klasifikasi <i>Naïve Bayes</i> .....	29
3.5	METODE PENGUMPULAN DATA .....	29
3.5.1	<i>Data Latih</i> .....	29
3.5.2	<i>Data Uji</i> .....	30
3.6	<i>PREPROCESSING</i> .....	31
3.6.1	<i>Case Folding</i> .....	32
3.6.2	<i>Cleansing</i> .....	33
3.6.3	<i>Convert Emoji</i> .....	33
3.6.4	<i>Convert Negation</i> .....	34
3.6.5	<i>Tokenizing</i> .....	34
3.6.6	<i>Normalization</i> .....	34
3.6.6.1	<i>Normalization kata typo</i> .....	35
3.6.6.2	<i>Normalization kata singkatan</i> .....	35
3.6.6.3	<i>Normalization kata slang</i> .....	36
3.6.6.4	<i>Normalization kata baku</i> .....	36
3.6.6.5	<i>Normalization kata bahasa inggris</i> .....	36
3.6.7	<i>Stopword Removal</i> .....	37
3.6.8	<i>Stemming</i> .....	37
3.7	PEMBOBOTAN TF-IDF .....	38
3.8	<i>NAÏVE BAYES CLASSIFICATION</i> .....	42
3.8.1	<i>Proses Latih</i> .....	43
3.8.2	<i>Proses Uji</i> .....	45
3.9	EVALUASI .....	50
BAB IV HASIL DAN PEMBAHASAN .....		55

4.1	IMPLEMENTASI.....	55
4.1.1	<i>Pengumpulan Data Latih</i> .....	55
4.1.2	<i>Pengumpulan Data Uji</i> .....	58
4.1.3	<i>Proses Preprocessing</i> .....	60
a)	<i>Case Folding</i> .....	60
b)	<i>Cleansing</i> .....	60
c)	<i>Convert Emoji</i> .....	61
d)	<i>Convert Negation</i> .....	62
e)	<i>Tokenizing</i> .....	63
f)	<i>Normalization</i> .....	63
g)	<i>Stopword Removal</i> .....	64
h)	<i>Stemming</i> .....	64
4.1.4	<i>Pembobotan TF-IDF</i> .....	65
4.1.5	<i>Pembobotan Naïve Bayes</i> .....	66
4.1.6	<i>Klasifikasi Naïve Bayes</i> .....	68
4.1.7	<i>Desain dan Implementasi GUI</i> .....	70
a).	Halaman Utama.....	70
b).	Halaman <i>import</i> data latih.....	71
c).	Halaman <i>preprocessing</i> data latih.....	72
d).	Halaman pembobotan TF-IDF.....	72
f).	Halaman <i>Crawling</i> data uji.....	74
g).	Halaman <i>preprocessing</i> data uji.....	74
h).	Halaman Klasifikasi <i>Naïve Bayes</i> .....	75
4.2	PENGUJIAN DAN PEMBAHASAN SISTEM.....	76
BAB V PENUTUP.....		78
5.1	KESIMPULAN.....	78
5.2	SARAN.....	80
DAFTAR PUSTAKA.....		81

## DAFTAR TABEL

<b>Tabel 2. 1.</b> Tabel Perbandingan dengan Penelitian Terdahulu .....	8
<b>Tabel 2. 2.</b> Tabel Confusion Matrix .....	18
<b>Tabel 3. 1.</b> Tabel Contoh Data Latih .....	30
<b>Tabel 3. 2.</b> Tabel Contoh Data Uji .....	31
<b>Tabel 3. 3.</b> Tabel Case Folding.....	32
<b>Tabel 3. 4.</b> Tabel Cleansing.....	33
<b>Tabel 3. 5.</b> Tabel Convert emoji.....	33
<b>Tabel 3. 6.</b> Tabel Convert Negation .....	34
<b>Tabel 3. 7.</b> Tabel Tokenizing.....	34
<b>Tabel 3. 8.</b> Tabel Normalization konversi kata typo .....	35
<b>Tabel 3. 9.</b> Tabel Normalization konversi kata singkatan .....	35
<b>Tabel 3. 10.</b> Tabel Normalization konversi kata slang.....	36
<b>Tabel 3. 11.</b> Tabel Normalization konversi kata baku.....	36
<b>Tabel 3. 12.</b> Tabel Normalization konversi kata inggris .....	37
<b>Tabel 3. 13.</b> Tabel Stopword Removal.....	37
<b>Tabel 3. 14.</b> Tabel stemming .....	38
<b>Tabel 3. 15.</b> Tabel koleksi data latih.....	38
<b>Tabel 3. 16.</b> Tabel koleksi dataset term.....	39
<b>Tabel 3. 17.</b> Tabel Contoh hasil TF-IDF pada dataset term .....	40
<b>Tabel 3. 18.</b> Tabel Contoh kasus Data latih .....	43
<b>Tabel 3. 19.</b> Tabel Perhitungan Probabilitas Data latih.....	44
<b>Tabel 3. 20.</b> Tabel Contoh kasus Data uji .....	45
<b>Tabel 3. 21.</b> Tabel Nilai Probabilitas Data uji.....	49
<b>Tabel 3. 22.</b> Confusion Matrix bobot minimum 0.....	51
<b>Tabel 3. 23.</b> Confusion Matrix bobot minimum 5.....	51
<b>Tabel 3. 24.</b> Confusion Matrix bobot minimum 10.....	52
<b>Tabel 3. 25.</b> Confusion Matrix bobot minimum 15.....	52
<b>Tabel 3. 26.</b> Confusion Matrix bobot minimum 20.....	53
<b>Tabel 3. 27.</b> Confusion Matrix bobot minimum 25.....	53
<b>Tabel 4. 1.</b> Tabel Komposisi Data Latih.....	57

<b>Tabel 4. 2.</b> Tabel Hasil Akurasi.....	76
<b>Tabel 5. 1.</b> Tabel Hasil Akurasi.....	79
<b>Tabel 5. 2</b> Tabel Acuan Bisnis Positif.....	80
<b>Tabel 5. 3</b> Tabel Acuan Bisnis Negatif .....	80

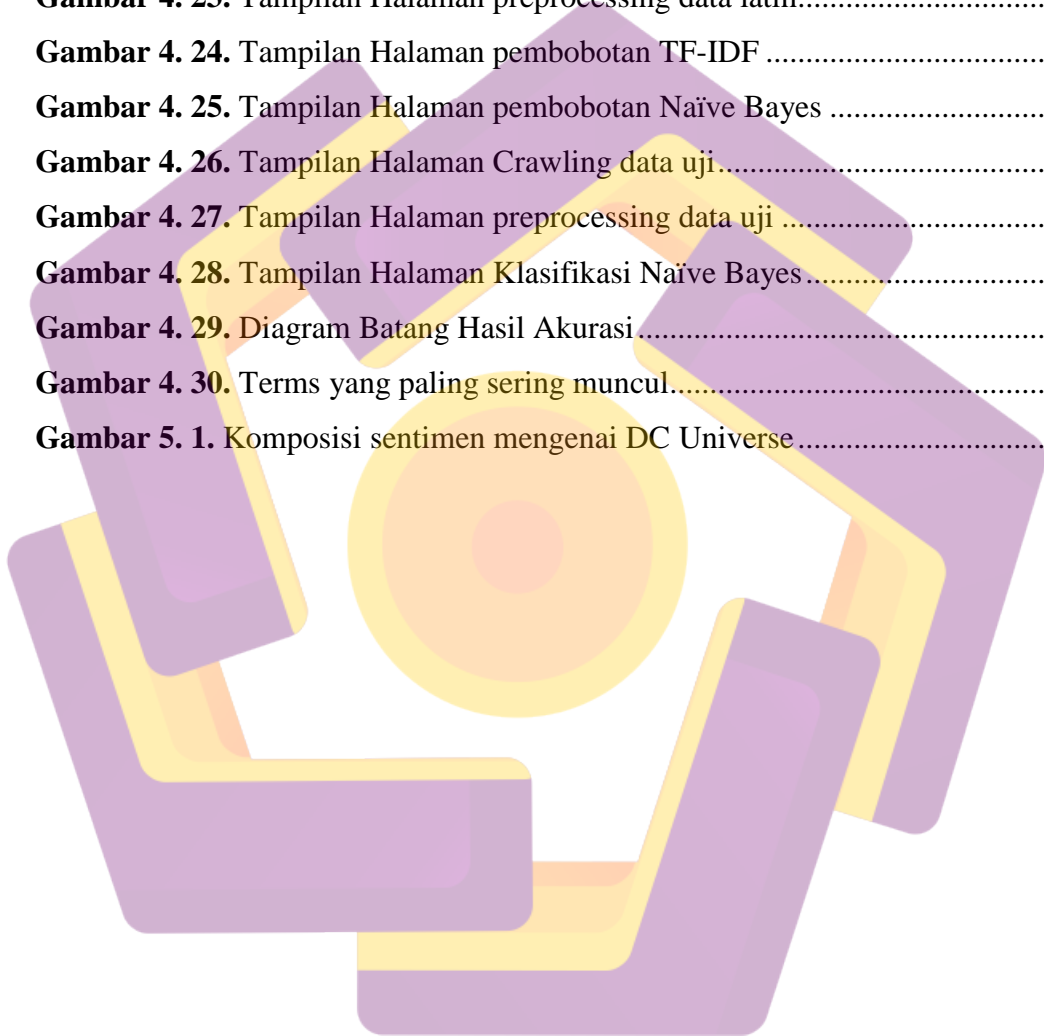


## DAFTAR GAMBAR

<b>Gambar 3. 1.</b> Alur Gambaran Umum .....	22
<b>Gambar 3. 2.</b> Alur Penelitian.....	24
<b>Gambar 3. 3.</b> Antar Muka Halama Utama.....	25
<b>Gambar 3. 4.</b> Antar Muka Import Data latih .....	26
<b>Gambar 3. 5.</b> Antar Muka Preprocessing Data latih.....	26
<b>Gambar 3. 6.</b> Antar Muka Proses Bobot TF-IDF .....	27
<b>Gambar 3. 7.</b> Antar Muka Proses Bobot Naïve Bayes .....	27
<b>Gambar 3. 8.</b> Antar Muka Crawling Data uji .....	28
<b>Gambar 3. 9.</b> Antar Muka Preprocessing Data uji.....	28
<b>Gambar 3. 10.</b> Antar Muka Klasifikasi Naïve Bayes .....	29
<b>Gambar 3. 11.</b> Diagram Alur Text Preprocessing .....	32
<b>Gambar 3. 12.</b> Diagram Alur Naïve Bayes Classifier .....	42
<b>Gambar 4. 1.</b> Command Proses Scraping .....	56
<b>Gambar 4. 2.</b> Hasil Proses Scraping .....	56
<b>Gambar 4. 3.</b> Data latih setelah dilabelkan manual .....	57
<b>Gambar 4. 4.</b> Proses import PHPExcel.....	58
<b>Gambar 4. 5.</b> Fungsi koneksi dengan Token .....	59
<b>Gambar 4. 6.</b> Fungsi crawling data uji .....	59
<b>Gambar 4. 7.</b> Koding proses case folding.....	60
<b>Gambar 4. 8.</b> Koding proses cleansing.....	61
<b>Gambar 4. 9.</b> Koding proses convert emoji .....	62
<b>Gambar 4. 10.</b> Koding proses convert negation .....	62
<b>Gambar 4. 11.</b> Koding proses tokenizing .....	63
<b>Gambar 4. 12.</b> Koding proses normalization.....	63
<b>Gambar 4. 13.</b> Koding proses Stopword Removal .....	64
<b>Gambar 4. 14.</b> Koding proses Stemming.....	64
<b>Gambar 4. 15.</b> Koding proses Pembobotan TF-IDF.....	65
<b>Gambar 4. 16.</b> Koding proses Probabilitas NB Positif.....	66
<b>Gambar 4. 17.</b> Koding proses Probabilitas NB Negatif .....	67



<b>Gambar 4. 18.</b> Koding proses Probabilitas NB Netral .....	68
<b>Gambar 4. 19.</b> Koding proses Klasifikasi Naïve Bayes Classifier .....	69
<b>Gambar 4. 20.</b> Tampilan Halaman Utama .....	70
<b>Gambar 4. 21.</b> Tampilan Halaman hasil import data latih.....	71
<b>Gambar 4. 22.</b> Tampilan Halaman form import data latih .....	71
<b>Gambar 4. 23.</b> Tampilan Halaman preprocessing data latih.....	72
<b>Gambar 4. 24.</b> Tampilan Halaman pembobotan TF-IDF .....	73
<b>Gambar 4. 25.</b> Tampilan Halaman pembobotan Naïve Bayes .....	73
<b>Gambar 4. 26.</b> Tampilan Halaman Crawling data uji.....	74
<b>Gambar 4. 27.</b> Tampilan Halaman preprocessing data uji .....	75
<b>Gambar 4. 28.</b> Tampilan Halaman Klasifikasi Naïve Bayes.....	75
<b>Gambar 4. 29.</b> Diagram Batang Hasil Akurasi.....	77
<b>Gambar 4. 30.</b> Terms yang paling sering muncul.....	77
<b>Gambar 5. 1.</b> Komposisi sentimen mengenai DC Universe.....	78



## INTISARI

DC Universe merupakan sebuah alam semesta fiksi dimana kumpulan pahlawan super dan penjahat super berdasarkan pada karakter yang muncul dalam buku-buku komik oleh DC Comics ada didalamnya. DC Comics sendiri adalah penerbit buku komik terbesar dan tertua di amerika milik DC Entertainment yang memproduksi dan menampilkan pahlawan super dan penjahat super.

Untuk memulai sebuah bisnis bertema pahlawan super, ada beberapa contoh bisnis yang bisa jadi referensi dan juga dapat dimanfaatkan untuk meraup keuntungan dari bisnis tersebut yaitu, persewaan komik, menjual *merchandise*, membuat *clothing line*, membuat kostum cosplay, membuat makanan bertema pahlawan super, menjual *action figure*. Dalam upaya untuk memulai bisnis bertema pahlawan super khususnya bertema DC Universe, perlu memperhatikan / mendengarkan konsumen DC Universe di Indonesia. Dengan mengetahui sentimen, pendapat, keinginan masyarakat Indonesia terhadap DC Universe, maka hal itu dapat digunakan sebagai acuan untuk memulai bisnis bertema pahlawan super di Indonesia.

Klasifikasi opini atau analisis sentimen adalah salah satu cara untuk mengetahui sentimen seseorang atau sekelompok orang terhadap produk, layanan, isu atau golongan tertentu dari berbagai *platform* media sosial dan *internet*. Pengguna media sosial bebas untuk mengekspresikan pendapat mereka. *Twitter* adalah salah satu media sosial yang disukai oleh masyarakat Indonesia. Penelitian ini mencoba untuk memanfaatkan apa yang ditulis oleh pengguna media sosial

*Twitter* atau lebih dikenal sebagai sebuah *tweet*. *Tweet* akan diproses dengan text mining dan diproses lagi menggunakan algoritma *Naïve Bayes Classifier*.

**Kata Kunci:** Sosial Media, Analisis Sentimen, *Naïve Bayes Classifier*



## **ABSTRACT**

*DC Universe is a fictional universe in which a collection of superheroes and super villains based on characters that appear in comic books by DC Comics is in it. DC Comics itself is the largest and oldest comic book publisher in America owned by DC Entertainment that produces and displays superheroes and super villains.*

*To start a super hero-themed business, there are a number of business examples that can be used as references and can also be used to reap profits from the business, namely, rental of comics, selling merchandise, making clothing lines, making cosplay costumes, making superhero-themed foods, selling action figure. In an effort to start a super hero-themed business, especially the DC Universe theme, it is necessary to pay attention / listen to DC Universe consumers in Indonesia. By knowing the sentiments, opinions, desires of the people of Indonesia towards DC Universe, then it can be used as a reference to start a super hero-themed business in Indonesia.*

*Classification of opinions or sentiment analysis is one way to find out about a person or group of people towards certain products, services, issues or groups from various social media platforms and the internet. Twitter is one of the social media that is loved by the people of Indonesia. This research tries to utilize what was written by Twitter social media users or better known as a tweet. Tweets will be processed by text mining and processed again using the Naïve Bayes Classifier algorithm.*

**Keyword:** *Social Media, Sentiment Analysis, Naïve Bayes Classifier*