

**ANALISIS OPINI PUBLIK TERHADAP FILM MARVEL STUDIOS PADA  
SOSIAL MEDIA TWITTER MENGGUNAKAN METODE NAIVE BAYES  
CLASSIFIER**

**SKRIPSI**



disusun oleh:

**Achmad Firdaus Olliv Pratama**

**16.11.0230**

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

**2022**

**ANALISIS OPINI PUBLIK TERHADAP FILM MARVEL STUDIOS PADA  
SOSIAL MEDIA TWITTER MENGGUNAKAN METODE NAIVE BAYES  
CLASSIFIER**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana pada Program Studi  
Informatika



disusun oleh:

**Achmad Firdaus Olliv Pratama**

**16.11.0230**

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

**2022**

**PERSETUJUAN**

**SKRIPSI**

**ANALISIS OPINI PUBLIK TERHADAP FILM MARVEL STUDIOS  
PADA SOSIAL MEDIA TWITTER MENGGUNAKAN METODE NAIVE  
BAYES CLASSIFIER**

yang dipersiapkan dan disusun oleh

**Achmad Firdaus Olliv Pratama**

**16.11.0230**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 13 September 2021

**Dosen Pembimbing,**

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

**NIK. 190302029**

# **PENGESAHAN**

## **SKRIPSI**

### **ANALISIS OPINI PUBLIK TERHADAP FILM MARVEL STUDIOS PADA SOSIAL MEDIA TWITTER MENGGUNAKAN METODE NAIVE BAYES CLASSIFIER**

yang dipersiapkan dan disusun oleh

**Achmad Firdaus Olliv Pratama**

**16.11.0230**

telah dipertahankan di depan Dewan Penguji

pada tanggal 21 Oktober 2021

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Erni Seniwati, S.kom.,M.Cs.**

**NIK. 190302231**

**Sri Ngudi Wahyuni, S.T.,M.kom.**

**NIK. 190302060**

**Bambang Sudaryanto, Drs., MM.**

**NIK. 190302029**

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

Tanggal 18 Januari 2022

**DEKAN FAKULTAS ILMU KOMPUTER**

**Hanif Al Fatta, M.Kom.**

**NIK. 190302096**

## PERNYATAAN

### 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, 18 Januari 2022



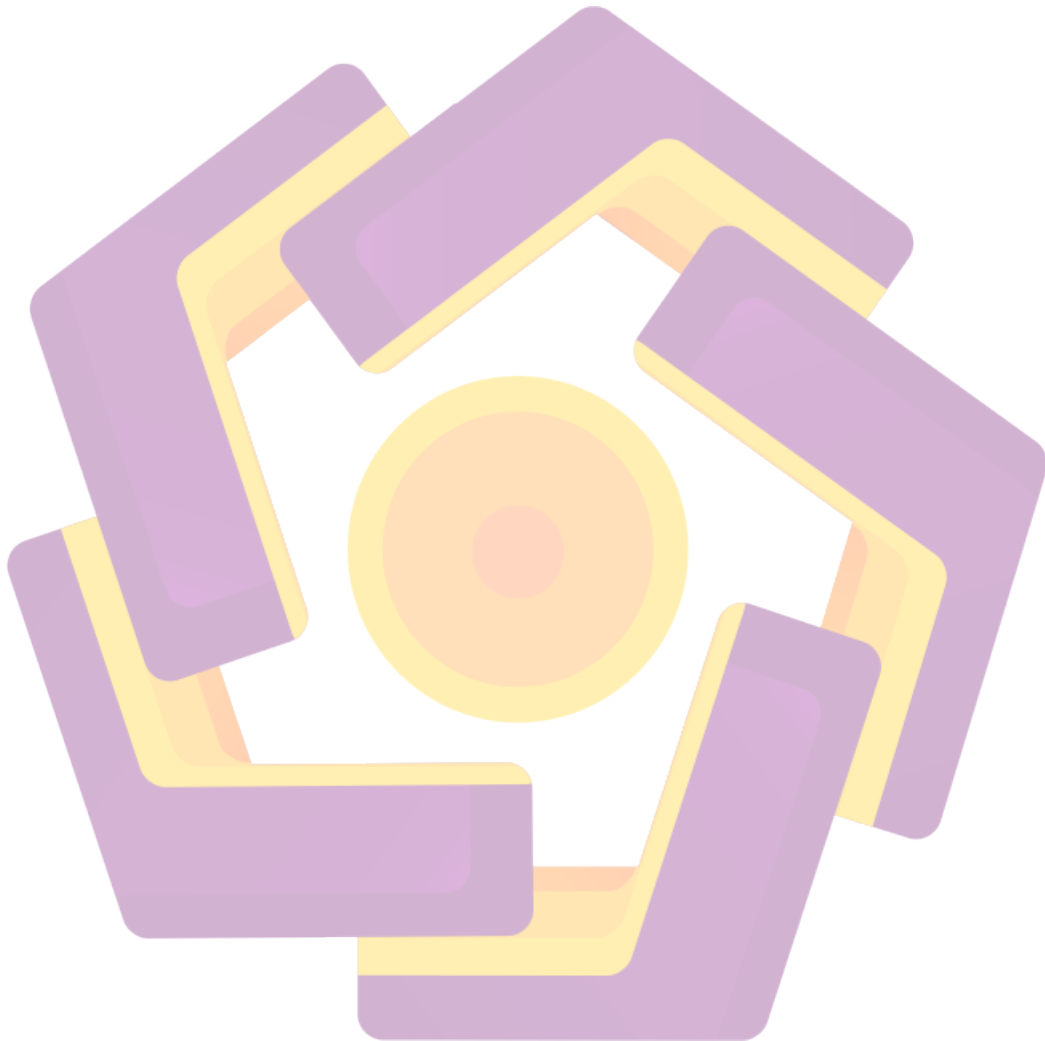
Achmad Firdaus Oliv Pratama

NIM. 16.11.0230

**MOTTO**

*“ To infinity, and beyond ”*

-Buzz Lightyear-



## PERSEMBAHAN

Alhamdulillahirobbil Alamin, puji syukur kehadiran Allah SWT yang telah memberikan rahmat, hidayah, kesehatan, kemudahan, serta kemampuan kepada saya, sehingga saya dapat menyelesaikan skripsi ini. Pada halaman persembahan ini, saya ingin berterimakasih yang sebesar-besarnya kepada:

1. Ibu Siti Alifah Nur'aini dan Bapak Sholeh selaku kedua orang tua saya yang telah memberikan segalanya kepada saya sehingga saya bisa berada di titik sekarang,
2. Bapak Bambang Sudaryanto, Drs., MM selaku dosen pembimbing saya yang senantiasa memberikan arahan dan motivasi kepada saya agar bisa menyelesaikan skripsi,
3. Andrew Patrick Defretes, Fahmi Ilmawan Sulaiman, Budi, Dwi selaku teman seperjuangan yang telah banyak membagikan motivasi dan materi sehingga saya bisa menyelesaikan skripsi sesuai target,
4. Untuk diri saya sendiri yang sudah percaya pada diri sendiri untuk melalui segala rintangan dalam menyelesaikan skripsi.

## KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT yang selalu melimpahkan rahmat dan karunia-nya kepada setiap hamba-nya dan tak lupa shalawat serta salam kepada junjungan Nabi besar kita, Nabi Muhammad SAW.

Skripsi ini disusun sebagai salah satu syarat dalam kelulusan pada Fakultas Ilmu Komputer Jurusan Sistem Informasi di Universitas Amikom Yogyakarta. Dalam penyusunan skripsi ini, berbagai pihak telah membantu penulis dalam segala hal. Sehingga penulis ingin mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. M. Suyanto, MM selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bambang Sudaryanto, Drs., MM selaku dosen pembimbing yang selalu memberikan arahan dan motivasi.
3. Bapak dan Ibu Dosen Universitas AMIKOM Yogyakarta yang telah memberikan banyak ilmu yang sangat bermanfaat bagi saya kedepannya.
4. Keluarga besar S1 Informatika yang telah menemani selamaperkuliahan dan memberikan kenangan yang luar biasa.

Penulis juga memohon maaf apabila dalam penyusunan skripsi ini masih banyak kekurangan dan masih jauh dari kata sempurna. Oleh karena itu penulis berharap kepada semua pihak agar dapat menyampaikan kritik dan saran yang membangun untuk menambah kesempurnaan skripsi ini.

Yogyakarta, 18 Januari 2022

Penulis



## DAFTAR ISI

JUDUL .....	i
PENGESAHAN .....	iii
PERNYATAAN.....	iv
MOTTO .....	v
PERSEMBAHAN.....	vi
KATA PENGANTAR .....	vii
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xii
DAFTAR GAMBAR .....	xiv
INTISARI.....	xvi
<i>ABSTRACT</i> .....	xvii
BAB I PENDAHULUAN .....	1
<b>1.1 Latar Belakang</b> .....	1
<b>1.2 Rumusan Masalah</b> .....	2
<b>1.3 Batasan Masalah</b> .....	2
<b>1.4 Maksud dan Tujuan Penelitian</b> .....	3
<b>1.5 Manfaat Penelitian</b> .....	4
<b>1.6 Metode Penelitian</b> .....	4
<b>1.6.1 Metode Pengumpulan Data</b> .....	4
<b>1.6.2 Metode Analisis Data</b> .....	5
<b>1.6.3 Metode Perancangan</b> .....	5
<b>1.6.4 Metode Pengujian</b> .....	6
<b>1.7 Sistematika Penulisan</b> .....	6
BAB II LANDASAN TEORI .....	8

<b>2.1</b>	<b>Tinjauan Pustaka .....</b>	<b>8</b>
<b>2.2</b>	<b>Penelitian Terdahulu.....</b>	<b>9</b>
<b>2.2</b>	<b>Data Mining .....</b>	<b>10</b>
<b>2.3</b>	<b>Text Mining.....</b>	<b>10</b>
<b>2.4</b>	<b>Analisis Sentimen .....</b>	<b>11</b>
<b>2.5</b>	<b>TF-IDF.....</b>	<b>11</b>
<b>2.6</b>	<b>Preprocessing .....</b>	<b>12</b>
<b>2.6.1</b>	<b><i>Case Folding</i>.....</b>	<b>13</b>
<b>2.6.2</b>	<b><i>Cleansing</i> .....</b>	<b>13</b>
<b>2.6.3</b>	<b><i>Convert Emoji</i>.....</b>	<b>14</b>
<b>2.6.4</b>	<b><i>Convert Negation</i> .....</b>	<b>14</b>
<b>2.6.5</b>	<b><i>Tokenizing</i> .....</b>	<b>14</b>
<b>2.6.6</b>	<b><i>Normalization</i> .....</b>	<b>14</b>
<b>2.6.7</b>	<b><i>Stopword Removal</i>.....</b>	<b>16</b>
<b>2.7</b>	<b><i>Sastrawi Stemmer</i> .....</b>	<b>16</b>
<b>2.8</b>	<b><i>Naïve Bayes Classifier</i> .....</b>	<b>16</b>
<b>2.9</b>	<b><i>Confusion Matrix</i> .....</b>	<b>18</b>
<b>2.10</b>	<b><i>Flowchart</i> .....</b>	<b>20</b>
<b>2.11</b>	<b><i>DFD (Data Flow Diagram)</i> .....</b>	<b>22</b>
<b>2.12</b>	<b><i>ERD (Entity Relationship Diagram)</i> .....</b>	<b>23</b>
<b>2.13</b>	<b><i>Waterfall</i>.....</b>	<b>23</b>
<b>2.14</b>	<b><i>Metode Testing</i> .....</b>	<b>25</b>
<b>2.14.1</b>	<b><i>Black Box Testing</i> .....</b>	<b>25</b>
<b>2.14.2</b>	<b><i>White Box Testing</i> .....</b>	<b>26</b>
<b>2.15</b>	<b><i>PHP (Hypertext Preprocessor)</i> .....</b>	<b>26</b>

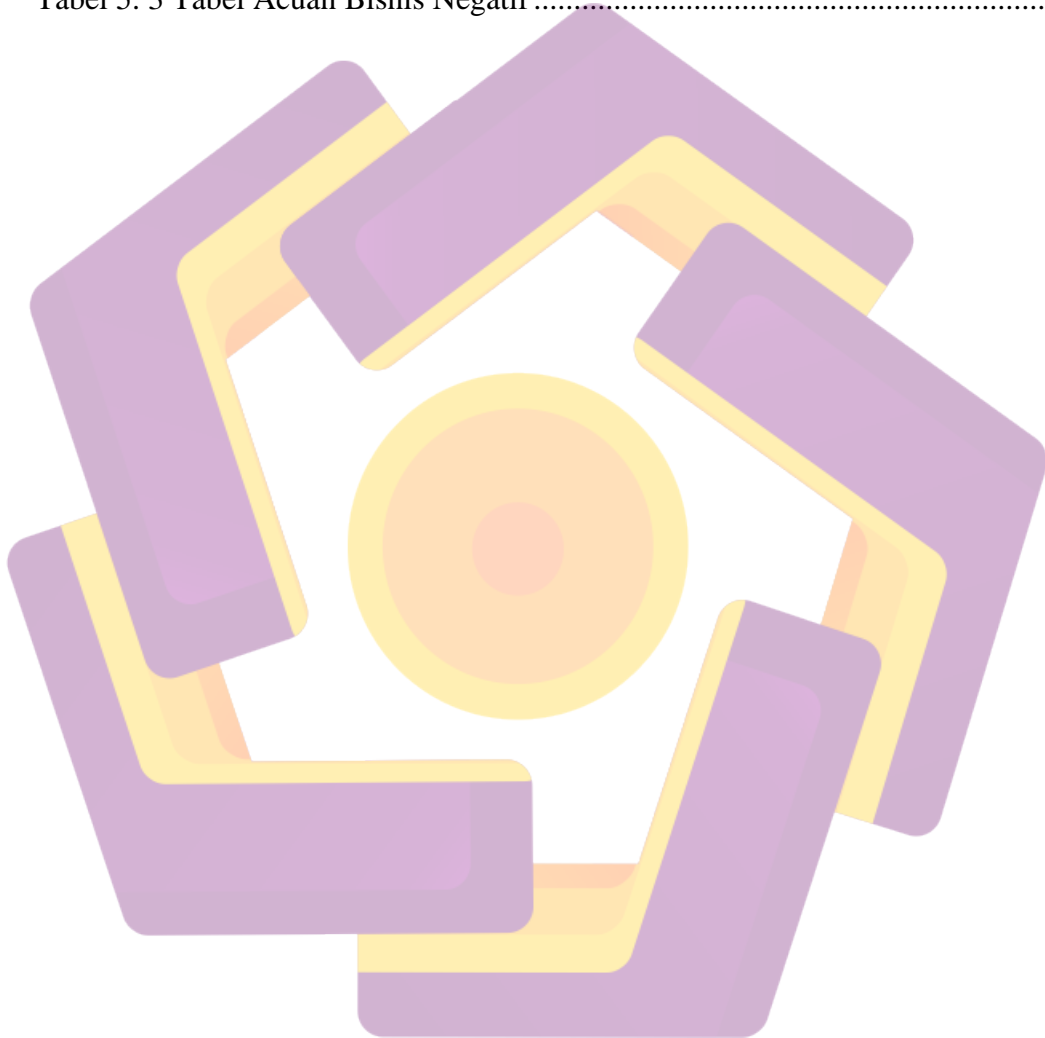
2.16	<i>Database</i> .....	27
2.17	<i>DBMS (Database Management System)</i> .....	27
2.18	<i>MySQL</i> .....	28
BAB III METODE PENELITIAN .....		29
3.1	<b>Gambaran Umum</b> .....	29
3.2	<b>Alat dan Bahan Penelitian</b> .....	30
3.2.1	<b>Alat penelitian</b> .....	30
3.2.2	<b>Bahan Penelitian</b> .....	31
3.3	<b>Alur Penelitian</b> .....	31
3.4	<b>Metode Pengumpulan Data</b> .....	33
3.4.1	<b>Data Training</b> .....	33
3.4.2	<b>Data Testing</b> .....	35
3.5	<b>Preprocessing</b> .....	35
3.5.1	<b>Case Folding</b> .....	36
3.5.2	<b>Cleansing</b> .....	37
3.5.3	<b>Convert Emoji</b> .....	37
3.5.4	<b>Convert Negation</b> .....	38
3.5.5	<b>Tokenzing</b> .....	38
3.5.6	<b>Normalization</b> .....	39
3.5.7	<b>Stop Watch Removal</b> .....	41
3.5.8	<b>Stemming</b> .....	42
3.6	<b>Pembobotan TF-IDF</b> .....	42
3.7	<b>Naïve Bayes Classification</b> .....	48
3.7.1	<b>Proses Training</b> .....	49
3.7.2	<b>Proses Testing</b> .....	52

3.8	<b>Evaluasi</b> .....	56
3.1	<b>Perancangan Sistem</b> .....	58
3.9.1	<i>Data Flow Diagram (DFD)</i> .....	58
3.1.1	<i>Perancangan Antar Muka</i> .....	60
BAB IV IMPLEMENTASI.....		67
4.1	<b>Implementasi</b> .....	67
4.1.1	<i>Pengumpulan Data Training</i> .....	67
4.1.2	<i>Pengumpulan Data Testing</i> .....	71
4.1.3	<i>Process Preprocessing</i> .....	73
4.1.4	<i>Pembobotan TF-IDF</i> .....	78
4.1.5	<i>Pembobotan Naïve Bayes</i> .....	78
4.1.6	<i>Klasifikasi Naïve Bayes</i> .....	81
4.1.7	<i>Desain dan Implementasi User Interface</i> .....	83
4.2	<b>Pengujian dan Pembahasan Sistem</b> .....	87
BAB V PENUTUP.....		90
4.1	<b>Kesimpulan</b> .....	90
5.2	<b>Saran</b> .....	91
DAFTAR PUSTAKA .....		95

## DAFTAR TABEL

Tabel 2. 1 Penelitian Sebelumnya.....	9
Tabel 2. 2 Confusion Matrix .....	19
Tabel 2.3 Simbol Flowchart.....	21
Tabel 2.4 Simbol-simbol DFD.....	22
Tabel 2.5 Simbol – Simbol ERD .....	23
Tabel 3. 1 Data Training .....	33
Tabel 3.2 Data Testing .....	35
Tabel 3.3 Case Folding .....	36
Tabel 3.4 Cleansing.....	37
Tabel 3.5 Convert Emoji.....	37
Tabel 3.6 Convert Negation .....	38
Tabel 3.7 Tokenzing.....	39
Tabel 3. 8 Normalization Typo .....	39
Tabel 3.9 Normalization Kata Singkatan .....	40
Tabel 3.10 Normalization Kata Slang .....	40
Tabel 3.11 Normalization Kata Baku.....	40
Tabel 3.12 Normalization Kata Bahasa Inggris .....	41
Tabel 3.13 Stop Watch Removal .....	41
Tabel 3.14 Stemming .....	42
Tabel 3.15 Pembobotan TF-IDF .....	42
Tabel 3.16 Terms/Kata.....	43
Tabel 3.17 Pembobotan TF-IDF .....	45
Tabel 3. 18 Proses Preprocessing.....	49
Tabel 3. 19 Probabilitas (PVj).....	50
Tabel 3. 20 Hasil Perhitungan Probabilitas .....	50
Tabel 3. 21 Proses Testing .....	52
Tabel 3.22 Hasil Probabilitas .....	55
Tabel 3.23 Confussion Matrix bobot minimum 0.....	56
Tabel 3.24 Confussion Matrix Bobot Minimum 3.....	57

Tabel 3.25 Confussion Matrix bobot minimum 5.....	57
Tabel 4. 1 Tabel Komposisi Data <i>Training</i> .....	70
Tabel 4. 2 Hasil Pengujian dan Pembahasan Sistem .....	87
Tabel 5. 1 Hasil Pembobotan TF-IDF.....	91
Tabel 5. 2 Tabel Acuan Bisnis Positif.....	91
Tabel 5. 3 Tabel Acuan Bisnis Negatif.....	91



## DAFTAR GAMBAR

Gambar 2.1 Alur Waterfall .....	24
Gambar 3.1 Alur Sistem.....	29
Gambar 3.2 Alur Penelitian.....	32
Gambar 3.3 Flowchart Preprocessing .....	36
Gambar 3. 4 Diagram Alur Naïve Bayes Classifier.....	48
Gambar 3.5 DFD Level 0.....	59
Gambar 3. 6 DFD Level 1 .....	60
Gambar 3.7 Halaman Dashboard .....	61
Gambar 3.8 Halaman Dataset Training.....	61
Gambar 3.9 Halaman Pembobotan Term.....	62
Gambar 3.10Halaman <i>Naïve Bayes Classifier</i> .....	63
Gambar 3.11 Halaman <i>Data Testing</i> .....	63
Gambar 3.12 Halaman Data Testing.....	64
Gambar 3.13 Halaman Data Testing.....	65
Gambar 4. 1 Command Proses Scraping Tweet.....	68
Gambar 4.2 Setting Untuk Menyimpan Hasil Scraping ke Database .....	68
Gambar 4.3 Hasil Simpan ke Database.....	69
Gambar 4.4 Hasil <i>Export File</i> ke .csv .....	70
Gambar 4.5 Proses <i>Import PHPExcel</i> .....	71
Gambar 4.6 Fungsi Koneksi Dengan Token.....	72
Gambar 4. 7 Fungsi <i>Crawling Data Uji</i> .....	72
Gambar 4. 8 Implementasi Untuk Case Folding.....	73
Gambar 4. 9 Implementasi Untuk Cleansing .....	74
Gambar 4. 10 Implementasi Proses Convert Emoji .....	75
Gambar 4. 11 Implementasi Proses Convert Negation .....	75
Gambar 4. 12 Implementasi Proses Tokenzing.....	76
Gambar 4. 13 Implementasi Proses Normalization.....	76
Gambar 4. 14 Impementasi Proses Stopword Removal.....	77
Gambar 4. 15 Impementasi Proses Stemming .....	77

Gambar 4. 16 Impementasi Proses TF-IDF .....	78
Gambar 4. 17 Impementasi Proses Probabilitas Positif .....	79
Gambar 4. 18 Impementasi Proses Probabilitas Neutral.....	80
Gambar 4. 19 Impementasi Proses Probabilitas Negatif.....	81
Gambar 4. 20 Implementasi Klasifikasi Naïve Bayes .....	82
Gambar 4. 21 Halaman Dashboard .....	83
Gambar 4. 22 Halaman Dataset Training.....	84
Gambar 4. 23 Halaman Pembobotan Term.....	84
Gambar 4. 24 Halaman Naïve Bayes Classifier.....	85
Gambar 4. 25 Halaman Naïve Bayes Classifier.....	85
Gambar 4. 26 Halaman Naïve Bayes Classifier.....	86
Gambar 4. 27 Halaman Classification Testing .....	87
Gambar 4. 28 Grafik Hasil Akurasi .....	88
Gambar 4. 29 10 Terms Yang Sering Muncul .....	89
Gambar 5.1 Diagram Pie Presentase Koleksi Data Training .....	90



## INTISARI

Marvel Studios merupakan studio film yang mungkin akan menjadi studio paling sukses dalam mengeluarkan film superhero dekade ini. Dari keseluruhan film yang telah dikeluarkan total ada 20 film hingga 2018. Hampir seluruh film yang dikeluarkan mendapat animo yang sangat besar dari masyarakat dunia, dan mendapat rating yang bagus juga.

Namun dari rating yang bagus tersebut tetap saja masih ada orang yang mengkritik baik dari kritikus profesional maupun akun pengkritik film pada social media, disini yang disorot lebih ke social media twitter karena menurut saya twitter tempat yang tepat untuk mendapat opini publik dikarenakan lebih bebas dari pada social media yang lainnya.

Tujuan dibuatnya analisis ini adalah untuk mengetahui pendapat public tentang film marvel studios yang lama maupun yang baru dari sudut pandang yang berbeda. Hingga akhirnya nanti dapat ditemukan berbagai pendapat bagaimana seharusnya film keluaran marvel studios yang akan ditayangkan seperti contohnya akankah film ini seharusnya mengikuti alur komiknya atau bisa ditambahkan jalan cerita yang lain sehingga banyak penggemar yang akan terpuaskan dengan film baru yang akan keluar.

**Kata Kunci:** *Marvel, Sentimen, Naïve Bayes, Twitter*

## **ABSTRACT**

*Marvel Studios is a film studio that might be the most successful studio in releasing superhero films this decade. Of the total films that have been issued a total of 20 films until 2018. almost all the films that were released received enormous interest from the world community and got a good rating too.*

*But from the good rating there are still people who criticize both professional critics and film critics on social media accounts, here are highlighted more on Twitter social media because in my opinion twitter the right place to get a public opinion is because it is freer than social other media.*

*The purpose of this analysis is to find out public opinion about old and new marvel studios from different perspectives. Until finally there can be found various opinions on how the film produced by Marvel Studios, which will be aired like for example, will the film should follow its commotion or another storyline so that many fans will be satisfied with the new film that will come out.*

**Keywords:** *Marvel, Sentimen, Naïve Bayes, Twitter*

