

**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI  
WEBTOON MENGGUNAKAN *TEXT MINING* DAN  
ALGORITMA SVM**

**SKRIPSI**

1. ARMELIA ISABELA TAEK : 20190040045
2. PUTRI ANUGRAH S 20190040096



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK KOMPUTER DAN DESAIN  
SUKABUMI  
JULI 2023**

**ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI  
WEBTOON MENGGUNAKAN *TEXT MINING* DAN**

**ALGORITMA SVM**

**SKRIPSI**

*Diajukan Untuk Memenuhi Salah Satu Syarat Dalam Menempuh  
Gelar Sarjana Teknik Informatika*

1. **Armelia Isabela Taek : 20190040045**
2. **Putri Anugrah S                      20190040096**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK KOMPUTER DAN DESAIN  
SUKABUMI  
JULI 2023**

## PERNYATAAN PENULIS

JUDUL : ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI  
WEBTOON MENGGUNAKAN *TEXT MINING* DAN  
ALGORITMA SVM

NAMA : ARMELIA ISABELA TAEK : 20190040045  
PUTRI ANUGRAH S : 20190040096

“Saya menyatakan dan bertanggungjawab dengan sebenarnya bahwa Skripsi ini adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya ada pihak lain yang mengklaim bahwa Skripsi ini sebagai karyanya, yang disertai dengan bukti-bukti yang cukup, maka saya bersedia untuk dibatalkan gelar Sarjana Teknik Informatika saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut”.

Sukabumi, 28 Juli 2023

  
Armelia Isabela Taek  
Penulis 1

  
Putri Anugrah S  
Penulis 2

## PERSETUJUAN SKRIPSI

JUDUL : ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI  
WEBTOON MENGGUNAKAN *TEXT MINING* DAN  
ALGORITMA SVM

NAMA : ARMELIA ISABELA TAEK : 20190040045  
PUTRI ANUGRAH S : 20190040096

Skripsi ini telah diperiksa dan disetujui

Sukabumi, 28 Juli 2023

Ketua Program Studi,

Pembimbing,



Anggun Fergina, M.Kom.

NIDN. 0407029301

Anggun Fergina, M.Kom.

NIDN. 0407029301

## PENGESAHAN SKRIPSI

JUDUL : ANALISIS SENTIMEN ULASAN PENGGUNA  
APLIKASI WEBTOON MENGGUNAKAN *TEXT*  
*MINING* DAN ALGORITMA SVM

NAMA PENULIS 1 : ARMELIA ISABELA TAEK : 20190040045

NAMA PENULIS 2 : PUTRI ANUGRAH S : 20190040096

Skripsi ini telah diujikan dan dipertahankan di depan Dewan Penguji pada Sidang Skripsi tanggal 11 Juli 2023. Menurut pandangan kami, Skripsi ini memadai dari segi kualitas untuk tujuan penganugerahan gelar Sarjana Komputer (S.Kom)

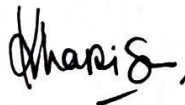
Sukabumi, 11 Juli 2023

Pembimbing I

Pembimbing II



Anggun Fergina, M.Kom.  
NIDN : 0407029301



Ivana Lucia Kharisma, M.Kom.  
NIDN : 0429038002

Ketua Dewan Penguji

Ketua Program Studi



Gina Purnama Insany, S.ST., M.Kom.  
NIDN : 0417077908



Anggun Fergina, M.Kom.  
NIDN : 0407029301

Dekan Fakultas Teknik, Komputer dan Desain

Ir. Paikun, S.T., M.T., IPM., Asean Eng

NIDN : 0402037401



*Skripsi ini kutujukan kepada  
Ayahanda dan Ibunda tercinta,  
dan Adik-adikku tersayang.  
Serta tak lupa kepada  
sahabat-sahabatku terkasih.  
~Be Brave & Be Humble~*

## ABSTRACT

*LINE Webtoon is a popular application for reading comics online, available on the Google Play Store. It has been around since 2015 and has a large user community. Out of a total of 35 million active LINE Webtoon users worldwide, approximately 6 million of them are from Indonesia. Despite the app's popularity and many 5-star reviews, not all users are satisfied with the performance of LINE Webtoon. This can be seen from user reviews in the comment section on the Play Store. New users often use these reviews as a reference to determine the best and most satisfying app to use. In this research, a sentiment analysis of LINE Webtoon users was conducted with the aim of classifying user reviews into positive and negative sentiments. The classification algorithm used was Support Vector Machine (SVM). The study collected a total of 15,000 data sets through web scraping from 2018 to 2023. The data was divided into three testing groups, with 90% for training and 10% for testing, 80% for training and 20% for testing, and the last one with 70% for training and 30% for testing. The results of testing with 90% training data and 10% testing data showed a precision score of 0.86, a recall score of 0.86, an f-measure score of 0.86, and an accuracy score of 0.82. The results of testing with 80% training data and 20% testing data showed a precision score of 0.86, a recall score of 0.85, an f-measure score of 0.85, and an accuracy score of 0.81. Lastly, the results of testing with 70% training data and 30% testing data showed a precision score of 0.85, a recall score of 0.85, an f-measure score of 0.85, and an accuracy score of 0.81. After analyzing the results of all three testing scenarios, the highest accuracy was achieved in the testing with 90% training data and 10% testing data, which was 0.82 or 82%.*

**Keywords :** Comment, Sentiment, Support Vector Machine, Webtoon



## ABSTRAK

LINE Webtoon adalah aplikasi populer untuk membaca komik secara *online* yang tersedia di Google Play Store. Aplikasi ini telah hadir sejak tahun 2015 dan memiliki komunitas pengguna yang besar. Dari total 35 juta pengguna aktif Line Webtoon di seluruh dunia, sekitar 6 juta di antaranya berasal dari Indonesia. Meskipun aplikasi ini sangat populer dan mendapatkan banyak ulasan bintang 5, tidak semua pengguna merasa puas dengan kinerja aplikasi LINE Webtoon. Hal ini terlihat dari ulasan yang diberikan oleh pengguna di kolom komentar di Play Store. Pengguna baru sering menggunakan ulasan pengguna sebagai acuan untuk menentukan aplikasi terbaik dan memuaskan untuk digunakan. Dalam penelitian ini, dilakukan analisis sentimen pengguna aplikasi Line Webtoon dengan tujuan mengklasifikasikan ulasan pengguna menjadi sentimen positif dan negatif. Algoritma yang digunakan untuk klasifikasi adalah *Support Vector Machine* (SVM). Penelitian ini menggunakan total 15.000 data set yang dikumpulkan melalui *web scraping* dari tahun 2018 sampai 2023. Data dibagi kedalam tiga kelompok pengujian, dimana data *training* 90% dan data *testing* 10%, data *training* 80% dan data *testing* 20%, dan yang terakhir data *training* 70% dan data *testing* 30%. Hasil dari pengujian data *training* 90% dan data *testing* 10% didapatkan nilai presisi 0.86, nilai *recall* 0.86, nilai *f-measure* 0.86 dan nilai akurasi 0.82. Hasil dari pengujian data *training* 80% dan data *testing* 20% didapatkan nilai presisi 0.86, nilai *recall* 0.85, nilai *f-measure* 0.85 dan nilai akurasi 0.81. Hasil dari pengujian data *training* 70% dan data *testing* 30% didapatkan nilai presisi 0.85, nilai *recall* 0.85, nilai *f-measure* 0.85 dan nilai akurasi 0.81. Maka setelah melihat hasil dari ketiga kali pengujian didapatkan nilai akurasi paling besar dari hasil pengujian data *training* 90% dan data *testing* 10% yaitu sebesar 0.82 atau 82%.

**Kata Kunci :** *Komentar, Sentimen, Support Vector Machine, Webtoon*

## KATA PENGANTAR

Puji syukur kami panjatkan ke hadirat Allah SWT, berkat rahmat dan karunia-Nya akhirnya penulis dapat menyelesaikan skripsi ini. Tujuan penulisan skripsi ini merupakan salah satu syarat untuk mencapai gelar Sarjana Komputer pada program Studi Teknik Informatika di Universitas Nusa Putra.

Sehubungan dengan itu penulis menyampaikan penghargaan dan ucapan terima kasih yang

sebesar-besarnya kepada :

1. Bapak Dr. Kurniawan ST, M.Si, MM Rektor Universitas Nusa Putra Sukabumi.
2. Bapak Anggy Pradiftha Junfithrana, S.Pd., MT selaku Wakil Rektor I Bidang Akademik Universitas Nusa Putra Sukabumi.
3. Ibu Anggun Fergina M.Kom selaku Kepala Program Studi Teknik Informatika Universitas Nusa Putra Sukabumi dan selaku dosen pembimbing I yang telah membimbing dan mengarahkan penulis dalam penulisan skripsi.
4. Ibu Ivana Lucia Kharisma M.Kom selaku Dosen Pembimbing II Universitas Nusa Putra Sukabumi yang telah membimbing dan mengarahkan penulis dalam penulisan skripsi.
5. Para Dosen Program Studi Teknik Informatika Universitas Nusa Putra Sukabumi yang telah berjasa memberikan ilmu pengetahuan.
6. Bapak Marsel Boja Taek dan Ibu Rianti sebagai orang tua dari Armelia Isabela Taek atas jasa, kesabaran, do'a dan tidak pernah lelah dalam mendidik dan memberi cinta yang tulus kepada penulis.
7. Bapak Dorateus Karmen Simanjorang dan Ibu Veronika Rusi Sinurat sebagai orang tua dari Putri Anugerah Simanjorang atas jasa, do'a dan tidak pernah lelah dalam mendidik dan memberi cinta yang tulus.
8. Adik-adik penulis yang tercinta Dedi Fernando Taek, Ardi Taek, Desi Novianti Taek, Yohanes Dian Putra Anugrah Simanjorang dan Gregorius Gorga Adolfo Anugrah Simanjorang.
9. Sahabat-sahabat penulis yang terkasih Siti Sa'diah, Sarah Sahira, Siti Syarah Sevia, Ismi Nurhadianti Rusmana, dan Kak Irwan yang telah menemani dan terus mendukung penulis untuk menyelesaikan skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, oleh karena itu kritik dan saran yang membangun dari berbagai pihak sangat kami harapkan demi perbaikan. Amin Yaa Rabbal 'Alamiin.

Sukabumi, Juli 2023

Penulis 1      Penulis 2

## HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik UNIVERSITAS NUSA PUTRA, saya yang bertanda tangan di bawah ini:

|               |                        |     |               |
|---------------|------------------------|-----|---------------|
| Nama          | : Armelia Isabela Taek | NIM | : 20190040045 |
| Nama          | : Putri Anugrah S      | NIM | : 20190040096 |
| Program Studi | : Teknik Informatika   |     |               |
| Jenis Karya   | : Skripsi              |     |               |

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Nusa Putra **Hak Bebas Royalti Noneksklusif (Non-exclusive Royalty- Free Right)** atas karya ilmiah saya yang berjudul:

**“ ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI WEBTOON MENGGUNAKAN *TEXT MINING* DAN ALGORITMA SVM”**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Noneksklusif ini Universitas Nusa Putra berhak menyimpan, mengalih media / format- kan, mengelola dalam bentuk pangkalan data (database), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di: Sukabumi

Pada tanggal: 28 Juli 2023

Yang menyatakan,

Mahasiswa

  
44AKX621729160  
**Armelia Isabela Taek**

Mahasiswa

  
A6AKX521729164  
**Putri Anugrah S**

## DAFTAR ISI

|                                                                                      |      |
|--------------------------------------------------------------------------------------|------|
| HALAMAN JUDUL .....                                                                  | i    |
| PERNYATAAN PENULIS .....                                                             | ii   |
| PERSETUJUAN SKRIPSI .....                                                            | iii  |
| PENGESAHAN SKRIPSI .....                                                             | iv   |
| ABSTRACT .....                                                                       | vi   |
| KATA PENGANTAR.....                                                                  | viii |
| HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK<br>KEPENTINGAN AKADEMIS ..... | ix   |
| DAFTAR ISI .....                                                                     | ix   |
| DAFTAR TABEL .....                                                                   | xiii |
| DAFTAR GAMBAR .....                                                                  | xivv |
| BAB I .....                                                                          | 1    |
| PENDAHULUAN.....                                                                     | 1    |
| 1.1 Latar Belakang .....                                                             | 1    |
| 1.2 Rumusan Masalah.....                                                             | 2    |
| 1.3 Batasan Masalah.....                                                             | 3    |
| 1.4 Tujuan Penelitian .....                                                          | 3    |
| 1.5 Manfaat Penelitian.....                                                          | 3    |
| 1.1.1 Manfaat bagi peneliti.....                                                     | 3    |
| 1.1.2 Manfaat bagi peneliti lain.....                                                | 4    |
| 1.1.3 Manfaat bagi masyarakat .....                                                  | 4    |
| 1.6 Sistematika Penulisan .....                                                      | 5    |
| BAB II.....                                                                          | 6    |
| TINJAUAN PUSTAKA.....                                                                | 6    |
| 2.1 Penelitian Terkait .....                                                         | 6    |
| 2.2 Landasan Teori.....                                                              | 11   |
| 2.2.1 LINE Webtoon .....                                                             | 11   |
| 2.2.2 Web Scraping .....                                                             | 12   |
| 2.2.3 Analisis Sentimen.....                                                         | 12   |
| 2.2.4 Text Mining .....                                                              | 13   |
| 2.2.5 Text Preprocessing.....                                                        | 14   |
| 2.2.6 Term Frequency Inverse Document Frequency (TF-IDF) .....                       | 15   |

|                                                                                       |    |
|---------------------------------------------------------------------------------------|----|
| 2.2.7 Metode <i>Sampling</i> ( <i>Oversampling</i> ) .....                            | 17 |
| 2.2.8 Algoritma <i>Support Vector Machine</i> (SVM) .....                             | 18 |
| 2.2.9 <i>Confusion Matrix</i> .....                                                   | 19 |
| 2.2.10 Web Browser.....                                                               | 20 |
| 2.2.11 Google Play Store.....                                                         | 20 |
| 2.2.12 Google Colab.....                                                              | 21 |
| 2.2.13 Python.....                                                                    | 21 |
| 2.2.14 Word cloud.....                                                                | 22 |
| 2.3 Kerangka Berpikir.....                                                            | 22 |
| BAB III.....                                                                          | 24 |
| METODOLOGI PENELITIAN .....                                                           | 24 |
| 3.1 Data Penelitian .....                                                             | 24 |
| 3.2 Pengolahan Data .....                                                             | 24 |
| 3.2.1 <i>Text Preprocessing</i> .....                                                 | 24 |
| 3.2.2 Pembobotan dengan TF-IDF .....                                                  | 27 |
| 3.2.3 Klasifikasi SVM.....                                                            | 27 |
| 3.2.4 <i>Oversampling</i> Data.....                                                   | 28 |
| 3.2.5 Evaluasi Model.....                                                             | 29 |
| 3.2.6 Visualisasi Data.....                                                           | 30 |
| 3.3 Diagram Alur Penelitian.....                                                      | 30 |
| 3.4 Implementasi Sistem.....                                                          | 31 |
| 3.4.1 Kebutuhan Sistem .....                                                          | 31 |
| 3.4.2 Perancangan Antar Muka .....                                                    | 32 |
| BAB IV .....                                                                          | 36 |
| HASIL DAN PEMBAHASAN .....                                                            | 36 |
| 4.1 Pengumpulan Data .....                                                            | 36 |
| 4.2 <i>Dataset</i> Mentah .....                                                       | 37 |
| 4.3 Pelabelan Sentimen Positif dan Negatif .....                                      | 38 |
| 4.4 Tahapan <i>Text-mining</i> .....                                                  | 39 |
| 4.4.1 <i>Text Processing</i> .....                                                    | 40 |
| 4.5 Pembobotan dengan <i>Term Frequency-Inverse Document Frequency</i> (TF-IDF) ..... | 46 |
| 4.6 <i>Oversampling</i> Data .....                                                    | 47 |
| 4.7 Klasifikasi <i>Support Vector Machine</i> (SVM).....                              | 48 |
| 4.8 Evaluasi Klasifikasi <i>Support Vector Machine</i> (SVM).....                     | 49 |

|                                       |    |
|---------------------------------------|----|
| 4.9 Visualisasi .....                 | 53 |
| 4.10 Analisis Deskriptif .....        | 55 |
| 4.11 Desain dan Implementasi GUI..... | 59 |
| BAB V .....                           | 71 |
| PENUTUP .....                         | 71 |
| 5.1 Kesimpulan .....                  | 71 |
| 5.2 Saran.....                        | 72 |
| DAFTAR PUSTAKA .....                  | 73 |



## DAFTAR TABEL

|                                                      | <b>Halaman</b> |
|------------------------------------------------------|----------------|
| Tabel 2.1 Penelitian Terkait .....                   | 6              |
| Tabel 3.1 Tabel Kontingensi .....                    | 32             |
| Tabel 4.1 <i>Dataset</i> Hasil <i>Stemming</i> ..... | 50             |
| Tabel 4.2 <i>Confusion Matrix</i> .....              | 50             |
| Tabel 4.3 Hasil Pengujian .....                      | 52             |



## DAFTAR GAMBAR

|                                                                                                  | Halaman |
|--------------------------------------------------------------------------------------------------|---------|
| Gambar 2.1 Ilustrasi Metode <i>Sampling</i> dan <i>Oversampling</i> .....                        | 17      |
| Gambar 2.2 Ilustrasi Model SVM .....                                                             | 18      |
| Gambar 2.3 <i>Confusion Matrix</i> .....                                                         | 19      |
| Gambar 2.4 Kerangka Berpikir .....                                                               | 23      |
| Gambar 3.1 <i>Flowchart Case Folding</i> .....                                                   | 25      |
| Gambar 3.2 <i>Flowchart Tokenization</i> .....                                                   | 25      |
| Gambar 3.3 <i>Flowchart Filtering</i> .....                                                      | 26      |
| Gambar 3.4 <i>Flowchart Stemming</i> .....                                                       | 27      |
| Gambar 3.5 Diagram Alir Penelitian.....                                                          | 30      |
| Gambar 3.6 Tampilan Halaman <i>Dashboard</i> .....                                               | 32      |
| Gambar 3.7 Tampilan Halaman <i>Dataset</i> .....                                                 | 33      |
| Gambar 3.8 Tampilan Halaman <i>Text Preprocessing</i> .....                                      | 33      |
| Gambar 3.9 Tampilan Data <i>Clean</i> .....                                                      | 33      |
| Gambar 3.10 Tampilan Halaman Data <i>Training</i> .....                                          | 34      |
| Gambar 3.11 Tampilan Halaman Data <i>Testing</i> .....                                           | 34      |
| Gambar 3.12 Tampilan Halaman Klasifikasi & Akurasi SVM.....                                      | 35      |
| Gambar 3.13 Tampilan Halaman Visualisasi Data .....                                              | 35      |
| Gambar 4.1 <i>Syntax Web Scraping</i> Google Play.....                                           | 37      |
| Gambar 4.2 <i>Dataset</i> Mentah Hasil <i>Scraping</i> .....                                     | 37      |
| Gambar 4.3 <i>Syntax</i> Pelabelan Positif dan Negatif .....                                     | 38      |
| Gambar 4.4 Hasil Pelabelan Kelas Sentimen .....                                                  | 39      |
| Gambar 4.5 <i>Syntax</i> Proses <i>Case Folding</i> .....                                        | 41      |
| Gambar 4.6 <i>Dataset</i> Hasil <i>Case Folding</i> .....                                        | 41      |
| Gambar 4.7 <i>Syntax</i> Proses <i>Tokenization</i> .....                                        | 42      |
| Gambar 4.8 <i>Dataset</i> Hasil <i>Tokenization</i> .....                                        | 43      |
| Gambar 4.9 <i>Syntax</i> Proses <i>Filtering</i> .....                                           | 44      |
| Gambar 4.10 <i>Dataset</i> Hasil <i>Filtering</i> .....                                          | 44      |
| Gambar 4.11 <i>Syntax</i> Proses <i>Stemming</i> .....                                           | 45      |
| Gambar 4.12 <i>Syntax</i> Proses Data TF-IDF .....                                               | 46      |
| Gambar 4.13 <i>Syntax</i> Proses <i>Oversampling</i> .....                                       | 47      |
| Gambar 4.14 <i>Dataset</i> Hasil <i>Oversampling</i> .....                                       | 48      |
| Gambar 4.15 <i>Syntax</i> Perhitungan Algoritma SVM.....                                         | 49      |
| Gambar 4.16 Hasil Prediksi dan Akurasi .....                                                     | 49      |
| Gambar 4.17 <i>Syntax</i> Perhitungan Akurasi, Presisi, <i>Recall</i> dan <i>F-measure</i> ..... | 50      |
| Gambar 4.18 Hasil Evaluasi Data <i>Testing</i> 10% .....                                         | 51      |
| Gambar 4.19 Hasil Evaluasi Data <i>Testing</i> 20% .....                                         | 51      |
| Gambar 4.20 Hasil Evaluasi Data <i>Testing</i> 30% .....                                         | 52      |
| Gambar 4.21 <i>Syntax</i> Visualisasi Diagram <i>Pie</i> .....                                   | 53      |
| Gambar 4.22 Diagram <i>Pie</i> Hasil Klasifikasi SVM .....                                       | 54      |
| Gambar 4.23 <i>Syntax</i> Visualisasi Wordcloud.....                                             | 54      |
| Gambar 4.24 Wordcloud Ulasan Aplikasi LINE Webtoon.....                                          | 55      |
| Gambar 4.25 Diagram Hasil Analisis Sentimen per Tahun .....                                      | 56      |
| Gambar 4.26 Diagram <i>Pie</i> Tahun 2018 .....                                                  | 56      |

|                                                                            |    |
|----------------------------------------------------------------------------|----|
| Gambar 4.27 Diagram <i>Pie</i> Tahun 2019 .....                            | 57 |
| Gambar 4.28 Diagram <i>Pie</i> Tahun 2020 .....                            | 57 |
| Gambar 4.29 Diagram <i>Pie</i> Tahun 2021 .....                            | 58 |
| Gambar 4.30 Diagram <i>Pie</i> Tahun 2022 .....                            | 58 |
| Gambar 4.31 Diagram <i>Pie</i> Tahun 2023 .....                            | 59 |
| Gambar 4.32 Tampilan Halaman <i>Dashboard</i> .....                        | 60 |
| Gambar 4.33 Tampilan Halaman <i>Dataset</i> .....                          | 60 |
| Gambar 4.34 Tampilan Halaman <i>Case Folding</i> .....                     | 61 |
| Gambar 4.35 Tampilan Halaman <i>Dataset Tokenization</i> .....             | 61 |
| Gambar 4.36 Tampilan Halaman <i>Dataset Filtering</i> .....                | 62 |
| Gambar 4.37 Tampilan Halaman <i>Dataset Stemming</i> .....                 | 62 |
| Gambar 4.38 Tampilan Halaman <i>Data Clean</i> .....                       | 63 |
| Gambar 4.39 Tampilan Halaman <i>Data Training 70%</i> .....                | 63 |
| Gambar 4.40 Tampilan Halaman <i>Data Training 80%</i> .....                | 64 |
| Gambar 4.41 Tampilan Halaman <i>Data Training 90%</i> .....                | 64 |
| Gambar 4.42 Tampilan Halaman <i>Data Testing 10%</i> .....                 | 65 |
| Gambar 4.43 Tampilan Halaman <i>Data Testing 20%</i> .....                 | 65 |
| Gambar 4.44 Tampilan Halaman <i>Data Testing 30%</i> .....                 | 66 |
| Gambar 4.45 Tampilan Halaman <i>Klasifikasi SVM Data Testing 10%</i> ..... | 67 |
| Gambar 4.46 Tampilan Halaman <i>Klasifikasi SVM Data Testing 20%</i> ..... | 68 |
| Gambar 4.47 Tampilan Halaman <i>Klasifikasi SVM Data Testing 30%</i> ..... | 69 |
| Gambar 4.48 Tampilan Halaman <i>Visualisasi Data</i> .....                 | 70 |



## BAB I

### PENDAHULUAN

#### 1.1 Latar Belakang

LINE Webtoon adalah salah satu aplikasi membaca komik secara *online* pada Google Play Store. Platform digital yang satu ini sangat populer di Indonesia. Aplikasi ini menawarkan berbagai macam komik dalam berbagai genre, seperti romantis, komedi, aksi, fantasi, dan horor. Aplikasi LINE Webtoon sendiri dapat diunduh dengan mudah di perangkat Android dan iOS. Di Indonesia LINE Webtoon, telah hadir sejak tahun 2015 dan memiliki komunitas pengguna yang besar. Dari total 35 juta pengguna aktif LINE Webtoon di seluruh dunia, sebanyak 6 juta pengguna aktif berasal dari Indonesia pada Agustus 2016. Hal ini menjadikan Indonesia sebagai pasar terbesar bagi LINE Webtoon [1]. Salah satu alasan utama kepopulerannya adalah aksesibilitasnya yang tinggi, karena dapat dibaca secara *online* melalui aplikasi di perangkat seluler. Meski kepopuleran aplikasi LINE Webtoon begitu besar dengan banyaknya bintang 5 yang didapat. Tak secara menyeluruh menggunakan memberikan ulasan yang positif, komentar negatif pun pasti tak luput didalamnya. Dalam konteks ini, unsur kepercayaan pengguna menjadi peranan yang penting untuk pengembangan dan bersaing dengan aplikasi komik *online* yang setiap saat banyak bermunculan. Kepercayaan pengguna atau masyarakat merupakan faktor kunci dalam membentuk persepsi mereka terhadap suatu produk atau layanan. Pendapat dan ulasan yang diberikan oleh pengguna dapat mempengaruhi pandangan dan keputusan orang lain. Oleh karena itu, penting untuk memahami berbagai sudut pandang dan pengalaman pengguna yang berbeda agar dapat membentuk gambaran yang lebih lengkap tentang aplikasi tersebut.

Akan tetapi dalam proses pelaksanaannya tak semudah itu. Ada banyak sekali ulasan komentar yang terus masuk di Google Play Store setiap waktunya, sulit bagi pengembang aplikasi jika harus secara manual memahami sentimen dari setiap ulasan komentar pengguna.

Maka dari itu diperlukan sebuah metode analisis sentimen yang efektif untuk mengklasifikasikan komentar pengguna menjadi sentimen positif dan negatif. Algoritma *Support Vector Machine* (SVM) telah terbukti menjadi salah satu metode yang efektif dalam melakukan klasifikasi teks berdasarkan sentimen. Dalam hal ini, analisis sentimen atau pendekatan *text mining* dengan algoritma SVM dapat membantu untuk memahami lebih lanjut ulasan pengguna, baik yang positif maupun negatif. Dengan menggunakan teknik-teknik tersebut, dapat dilakukan pengecekan dan pengelompokan ulasan berdasarkan sentimen yang terkandung di dalamnya, sehingga dapat memberikan wawasan yang lebih mendalam mengenai pandangan pengguna terhadap aplikasi LINE Webtoon.

Melalui penelitian ini, diharapkan dapat memberikan kontribusi dalam memahami sentimen pengguna terhadap aplikasi LINE Webtoon. Dengan menganalisis ulasan komentar pengguna menggunakan algoritma SVM, penelitian ini akan memberikan wawasan tentang bagaimana pengguna merespon dan merasakan aplikasi LINE Webtoon. Hasil penelitianpun diharapkan dapat membantu pengembang aplikasi dalam meningkatkan kualitas, fitur, konten, dan pengalaman pengguna di platform LINE Webtoon. Dengan memahami sentimen pengguna terhadap aplikasi LINE Webtoon, penelitian ini dapat memberikan wawasan tentang preferensi pengguna, kepuasan pengguna, dan potensi pembaruan dalam pengembangan aplikasi serupa di masa yang akan datang.

## 1.2 Rumusan Masalah

Rumusan masalah pada penelitian meliputi beberapa hal, diantaranya :

1. Bagaimana cara penerapan algoritma SVM untuk memprediksi ulasan pengguna aplikasi LINE Webtoon?
2. Apakah algoritma SVM mampu mengklasifikasikan sentimen pengguna pada ulasan komentar di aplikasi LINE Webtoon menjadi positif dan negatif secara akurat?

### 1.3 Batasan Masalah

Batasan masalah pada penelitian kali ini telah ditetapkan agar tidak menyimpang dari pokok pembahasan dan demi terciptanya tujuan penelitian yang telah direncanakan, adapun batasan masalah pada penelitian kali ini diantaranya :

1. Penelitian ini akan membatasi analisis sentimen terhadap ulasan komentar pengguna yang terkait dengan aplikasi LINE Webtoon pada platform Google Play Store.
2. Data ulasan komentar yang digunakan ditarik dari sumber yang terbatas, seperti Google Play Store, dengan jumlah data yang telah ditentukan sebelumnya.
3. Penelitian ini akan fokus pada analisis sentimen untuk mengklasifikasikan komentar-komentar pengguna menjadi sentimen positif atau negatif. Aspek-aspek lain dari sentimen, seperti sentimen netral atau sentimen berdasarkan kategori tertentu, tidak akan dibahas secara mendalam.
4. Metode yang akan digunakan dalam penelitian ini terbatas pada algoritma SVM sebagai metode utama untuk analisis sentimen.



### 1.4 Tujuan Penelitian

Adapun penelitian ini bertujuan sebagai berikut :

1. Untuk mengklasifikasikan ulasan komentar pengguna aplikasi LINE Webtoon pada Google Play Store dengan menggunakan algoritma *Support Vector Machine* (SVM) serta menghitung keakurasiannya.
2. Untuk menganalisis performa algoritma SVM dalam melakukan analisis sentimen pada ulasan komentar pengguna di aplikasi LINE Webtoon, dengan tujuan memahami sejauh mana algoritma tersebut efektif dalam mengklasifikasikan sentimen positif dan negatif.

### 1.5 Manfaat Penelitian

Berikut adalah beberapa manfaat dari penelitian ini:

#### 1.1.1 Manfaat bagi peneliti:

- a) Menambah pemahaman tentang analisis sentimen dan penerapannya dalam konteks aplikasi LINE Webtoon.
- b) Peneliti dapat memperluas pengetahuan dalam bidang ini dan mengembangkan metode atau teknik yang lebih baik dalam analisis sentimen.
- c) Memberikan dasar penelitian dan referensi untuk penelitian lanjutan yang berkaitan dengan analisis sentimen aplikasi LINE Webtoon atau penggunaan algoritma SVM dalam analisis sentimen pada domain lain.

#### 1.1.2 Manfaat bagi peneliti lain:

- a) Memperkaya literatur ilmiah di bidang analisis sentimen dan aplikasi LINE Webtoon. Penelitian ini dapat menjadi referensi bagi peneliti lain yang tertarik untuk melanjutkan penelitian dalam domain yang serupa atau berkaitan.
- b) Membuka peluang kolaborasi antara peneliti yang tertarik pada analisis sentimen, pengembangan aplikasi, dan pengalaman pengguna. Peneliti lain dapat membangun dan memperluas penelitian ini dengan menambahkan variabel atau aspek tambahan untuk memperoleh pemahaman yang lebih komprehensif.

#### 1.1.3 Manfaat bagi masyarakat

- a) Memberikan informasi yang berguna bagi pengguna aplikasi LINE Webtoon dalam memahami sentimen pengguna lain terhadap aplikasi tersebut. Hal ini dapat membantu mereka dalam membuat keputusan tentang menggunakan atau memilih aplikasi LINE Webtoon yang sesuai dengan preferensi mereka.
- b) Meningkatkan kualitas dan pengalaman pengguna di platform LINE Webtoon dengan memahami faktor-faktor yang mempengaruhi sentimen positif atau negatif pengguna. Pengembang dapat menggunakan hasil penelitian ini untuk melakukan perbaikan dan peningkatan fitur, konten, atau layanan yang disediakan dalam aplikasi LINE Webtoon.

### **1.6 Sistematika Penulisan**

Untuk memahami lebih jelas tentang penelitian ini, dilakukan pengelompokan materi menjadi beberapa sub bab dengan sistematika penulisan sebagai berikut:

#### **BAB I Pendahuluan**

Bab ini berisikan tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

#### **BAB II Landasan Penelitian**

Bab ini berisikan tentang penelitian terkait/terdahulu, landasan teori yang digunakan sebagai bahan dasar referensi dalam proses penelitian, dan kerangka berpikir.

#### **BAB III Metodologi Penelitian**

Bab ini berisikan tentang penjelasan dari metode yang digunakan dalam penelitian, sumber data dan jenis data serta metode analisis data.

#### **BAB IV Hasil dan Pembahasan**

Bab ini berisikan hasil-hasil tahapan penelitian, mulai dari analisis, hasil pengujian dan pembahasannya.

#### **BAB V Penutup**

Bab ini berisikan kesimpulan dan saran dari peneliti yang sifatnya membangun untuk penelitian selanjutnya.

#### **BAB VI Daftar Pustaka**

Berisi tentang sumber-sumber ataupun buku-buku yang menjadi referensi dalam pembuatan skripsi ini.



## BAB V

### PENUTUP

#### 5.1 Kesimpulan

Penelitian ini bertujuan untuk menganalisis sentimen dari ulasan komentar pengguna aplikasi LINE Webtoon menggunakan algoritma SVM. Data yang digunakan dalam penelitian ini terdiri dari 15.000 ulasan komentar yang ditarik dari tahun 2018 hingga 2023 melalui proses *scraping*. Dalam penelitian ini, ada beberapa kesimpulan yang dapat diambil, diantaranya :

- Dari hasil 3 kali percobaan yang dilakukan dengan data *training* 70%, 80%, 90% dan data *testing* 10%, 20%, 30% didapatkan hasil akurasi yang paling tinggi dari hasil pengujian data *training* 90% dan data *testing* 10%. Dengan hasil akurasi yang bernilai 0.82 atau 82%.
- Berdasarkan hasil akurasi dari ketiga pengujian dapat disimpulkan bahwa analisis sentimen masyarakat terhadap aplikasi LINE Webtoon cenderung positif. Dan terbukti bahwa algoritma *Support Vector Machine* (SVM) efektif untuk analisis sentimen pada ulasan komentar aplikasi LINE Webtoon. Dengan akurasi yang relatif tinggi, penggunaan algoritma SVM dalam penelitian ini dapat membantu mengidentifikasi sentimen pengguna secara otomatis dalam skala besar.
- Bahasa pemrograman yang digunakan dalam penelitian ini adalah PHP dan HTML. *Tools* yang digunakan untuk menarik data ulasan pengguna aplikasi LINE Webtoon ialah web scraping, *tools* yang digunakan untuk melakukan penelitian adalah Google Colab dengan *text preprocessing* yang dilakukan yaitu *case folding*, *tokenizing*, *filtering* dan *stemming*.
- Dari data pertahun yang didapatkan terlihat bahwa secara keseluruhan pengguna aplikasi LINE Webtoon mengalami peningkatan terhadap ulasan positif. Pada tahun 2020 menjadi tahun paling banyak mendapatkan ulasan positif. Dan secara berkala ulasan pengguna terus menunjukkan angka yang positif.

## 5.2 Saran

Terdapat beberapa saran untuk peneliti selanjutnya. Pertama, dapat melakukan pengujian dengan algoritma lain seperti *Naive Bayes*, *Random Forest*, atau *Decision Tree* untuk meningkatkan akurasi yang lebih optimal. Kedua, diharapkan dapat memperluas jangkauan data yang digunakan dengan menggabungkan ulasan komentar dari platform lain seperti Instagram atau Twitter. Terakhir peneliti selanjutnya dapat memasukkan data kamus bahasa gaul pada proses *stopword* untuk mengakomodasi penggunaan kata kurang baku terhadap ulasan komentar. Hal ini diperlukan karena masih banyak penggunaan bahasa-bahasa yang kurang baku dalam komentar tersebut. Dengan memasukkan kamus bahasa gaul, penelitian dapat mengidentifikasi kata atau ungkapan bahasa gaul yang lebih spesifik, sehingga analisis sentimen menjadi lebih akurat dan lebih sesuai dengan bahasa yang digunakan oleh pengguna di Google Play Store.



## DAFTAR PUSTAKA

- [1] A. F. Lestari, "LINE WEBTOON SEBAGAI INDUSTRI KOMIK DIGITAL", *Jurnal Ilmu Komunikasi*, vol. 6, no. 2, Oct. 2020.
- [2] F. F. Irfani, "ANALISIS SENTIMEN REVIEW APLIKASI RUANGGURU MENGGUNAKAN ALGORITMA SUPPORT VECTOR MACHINE," *JBMI (Jurnal Bisnis, Manajemen, dan Informatika)*, vol. 16, no. 3, pp. 258–266, Feb. 2020, doi: 10.26487/jbmi.v16i3.8607.
- [3] A. Muhammadin and I. A. Sobari, "Analisis Sentimen Pada Ulasan Aplikasi Kredivo Dengan Algoritma SVM Dan NBC," *Jurnal Rekayasa Perangkat Lunak*, vol. 2, no. 2, 2021, [Online]. Available: <http://jurnal.bsi.ac.id/index.php/reputasi>
- [4] R. Apriani *et al.*, "ANALISIS SENTIMEN DENGAN NAÏVE BAYES TERHADAP KOMENTAR APLIKASI TOKOPEDIA," 2019.
- [5] S. I. Nurhafida and F. Sembiring, "Analisis Sentimen Aplikasi Novel Online Di Google Play Store Menggunakan Algoritma Support Vector Machine (SVM)," 2022.
- [6] S. J and K. U, "Sentiment analysis of amazon user reviews using a hybrid approach," *Measurement: Sensors*, p. 100790, May 2023, doi: 10.1016/j.measen.2023.100790.
- [7] S. Fide, "ANALISIS SENTIMEN ULASAN APLIKASI TIKTOK DI GOOGLE PLAY MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM) DAN ASOSIASI," vol. 10, no. 3, pp. 346–358, 2021, [Online]. Available: <https://ejournal3.undip.ac.id/index.php/gaussian/>
- [8] M. Afdal, L. Rahma Elita, P. Studi Sistem Informasi, F. H. Sains dan Teknologi UIN Suska Riau, M. Soebrantas KM, and P. Pekanbaru -Riau, "PENERAPAN TEKNIK MINING PADA APLIKASI TOKOPEDIA MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 8, no. 1, 2022.
- [9] T. Elizabeth, "Analisis Sentimen Ulasan Aplikasi PrimaKu Menggunakan Metode Support Vector Machine," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 9, no. 4, 2022, [Online]. Available: <http://jurnal.mdp.ac.id>
- [10] M. Kurnia Maulidina and E. Itje Sela, . "ANALISIS SENTIMEN KOMENTAR WARGANET TERHADAP POSTINGAN INSTAGRAM MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER DAN TF-IDF (Studi Kasus: Instagram Gubernur Jawa Barat Ridwan Kamil)." 2020.
- [11] A. Indrawati, "PENERAPAN TEKNIK KOMBINASI OVERSAMPLING DAN UNDERSAMPLING UNTUK MENGATASI PERMASALAHAN IMBALANCED DATASET," *Jurnal Informatika dan Komputer) Akreditasi KEMENRISTEKDIKTI*, vol. 4, no. 1, 2021, doi: 10.33387/jiko.

- [12] M. Tingkatkemanisan, M. Berdasarkan, and F. Warna, "Klasifikasi Support Vector Machine (SVM) Untuk," *MIND Journal | ISSN*, vol. 3, no. 2, pp. 16–24, 2018, doi: 10.26760/mindjournal.
- [13] A. Algoritma, K. Pada, S. Rapidminer, and W. Ainurrohmah, "Akurasi Algoritma Klasifikasi pada Software Rapidminer dan Weka," *Prosiding Seminar Nasional Matematika*, vol. 4, pp. 493–499, 2021, [Online]. Available: <https://journal.unnes.ac.id/sju/index.php/prisma/>
- [14] I. Journal ----of , T. Rochmadi, and J. Brawijaya No, "'Tri Rochmadi' LIVE FORENSIK UNTUK ANALISA ANTI FORENSIK PADA WEB BROWSER STUDI KASUS BROWZAR," 2018. [Online]. Available: <https://ejournal.almaata.ac.id/index.php/IJUBI>
- [15] N. Herlinawati *et al.*, "ANALISIS SENTIMEN ZOOM CLOUD MEETINGS DI PLAY STORE MENGGUNAKAN NAÏVE BAYES DAN SUPPORT VECTOR MACHINE," 2020.
- [16] R. Gelar Guntara, "Pemanfaatan Google Colab Untuk Aplikasi Pendeteksian Masker Wajah Menggunakan Algoritma Deep Learning YOLOv7," *Jurnal Teknologi Dan Sistem Informasi Bisnis*, vol. 5, no. 1, pp. 55–60, Feb. 2023, doi: 10.47233/jteksis.v5i1.750.
- [17] R. M. Clinton and R. Sengkey, "Purwarupa Sistem Daftar Pelanggaran Lalulintas Berbasis Mini-Komputer Raspberry Pi," *Jurnal Teknik Elektro dan Komputer*, vol. 8, no. 3, pp. 181-192, 2019.
- [18] M. Galih Pradana, "PENGUNAAN FITUR WORDCLOUD DAN DOCUMENT TERM MATRIX DALAM TEXT MINING." 2020.

