

**ANALISIS SENTIMEN DENGAN ALGORITMA SUPPORT  
VECTOR MACHINE TERHADAP ULASAN APLIKASI DISNEY+  
HOTSTAR DI PLAY STORE**

**SKRIPSI**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**HASYMI AL HAMDI**

**19.11.2997**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2026**

**ANALISIS SENTIMEN DENGAN ALGORITMA SUPPORT  
VECTOR MACHINE TERHADAP ULASAN APLIKASI DISNEY+  
HOTSTAR DI PLAY STORE**

**SKRIPSI**

untuk memenuhi salah satu syarat mencapai derajat Sarjana  
Program Studi Informatika



disusun oleh

**HASYMI AL HAMDI**

**19.11.2997**

Kepada

**FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA**

**2026**

**HALAMAN PERSETUJUAN**

**SKRIPSI**

**ANALISIS SENTIMEN DENGAN ALGORITMA SUPPORT VECTOR  
MACHINE TERHADAP ULASAN APLIKASI DISNEY+ HOTSTAR DI  
PLAY STORE**

yang disusun dan diajukan oleh

**Hasymi Al Hamdi**

**19.11.2997**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 27 Februari 2026

**Dosen Pembimbing,**



**Windha Mega Pradnya Dhuhita, S.Kom., M.Kom.**

**NIK. 190302185**

**HALAMAN PENGESAHAN**

**SKRIPSI**

**ANALISIS SENTIMEN DENGAN ALGORITMA SUPPORT VECTOR  
MACHINE TERHADAP ULASAN APLIKASI DISNEY+ HOTSTAR DI  
PLAY STORE**

yang disusun dan diajukan oleh

**Hasymi Al Hamdi**

**19.11.2997**

Telah dipertahankan di depan Dewan Penguji  
pada tanggal 27 Februari 2026

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Dr. Hartatik, S.T., M.Cs.**  
**NIK. 190302232**

**Novi Prisma Yunita, M.Kom.**  
**NIK. 190302526**

**Windha Mega Pradnya Dhuhiita, S.Kom., M.Kom.**  
**NIK. 190302185**



Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 27 Februari 2026

**DEKAN FAKULTAS ILMU KOMPUTER**



**Prof. Dr. Kusrini, M.Kom.**  
**NIK. 190302106**

## HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Hasymi Al Hamdi  
NIM : 19.11.2997

Menyatakan bahwa Skripsi dengan judul berikut:

**ANALISIS SENTIMEN DENGAN ALGORITMA SUPPORT VECTOR MACHINE TERHADAP ULASAN APLIKASI DISNEY+ HOTSTAR DI PLAY STORE**

Dosen Pembimbing : Windha Mega Pradnya Duhita, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 24 Februari 2026

Yang Menyatakan,

  
7AANX207968960

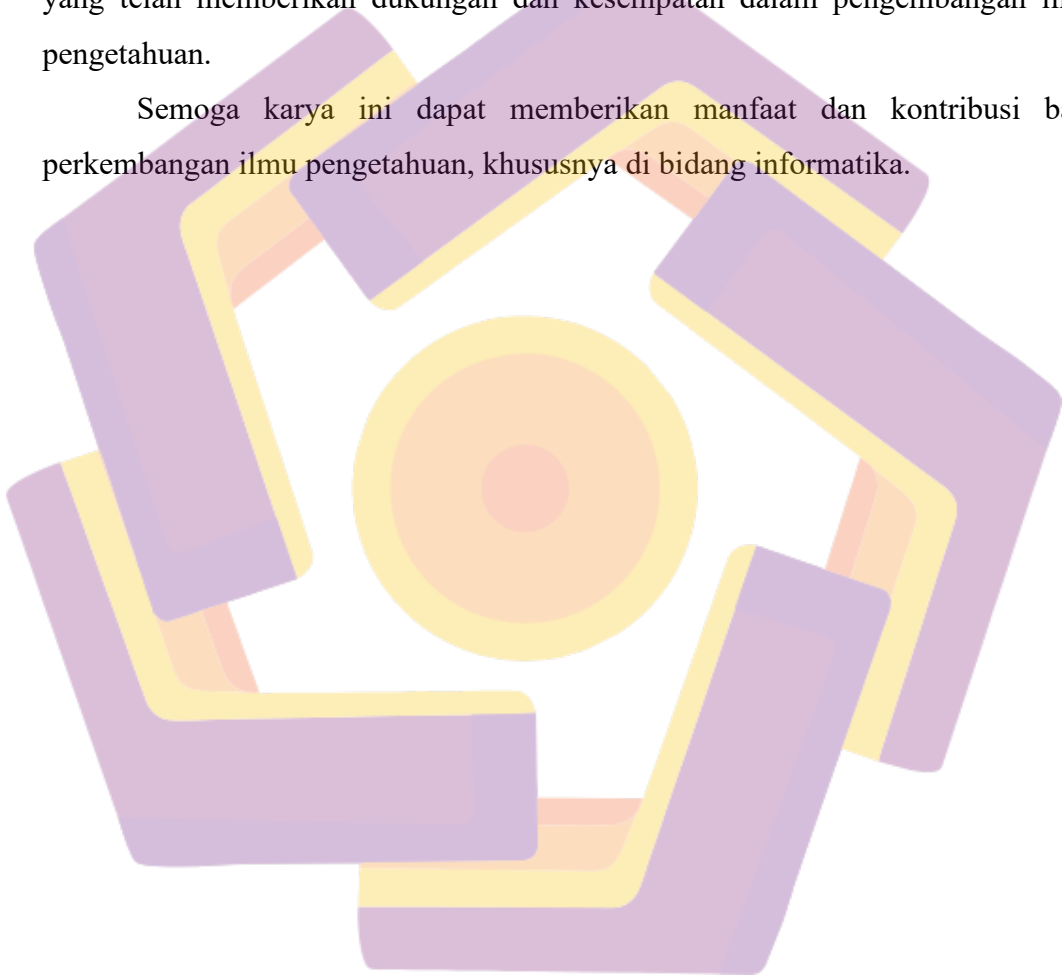
Hasymi Al Hamdi

## HALAMAN PERSEMBAHAN

Skripsi ini dipersembahkan sebagai wujud tanggung jawab akademik dan dedikasi penulis dalam menempuh pendidikan pada Program Studi S1 Informatika.

Karya ilmiah ini dipersembahkan kepada pihak-pihak yang berperan dalam proses pendidikan penulis, khususnya keluarga, dosen, serta lingkungan akademik yang telah memberikan dukungan dan kesempatan dalam pengembangan ilmu pengetahuan.

Semoga karya ini dapat memberikan manfaat dan kontribusi bagi perkembangan ilmu pengetahuan, khususnya di bidang informatika.



## KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga skripsi yang berjudul “ANALISIS SENTIMEN DENGAN ALGORITMA SUPPORT VECTOR MACHINE TERHADAP ULASAN APLIKASI DISNEY+ HOTSTAR DI PLAY STORE” dapat diselesaikan dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi S1 Informatika, Universitas AMIKOM Yogyakarta.

Dalam proses penyusunan skripsi ini, penulis memperoleh banyak bantuan, dukungan, serta bimbingan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan terima kasih kepada:

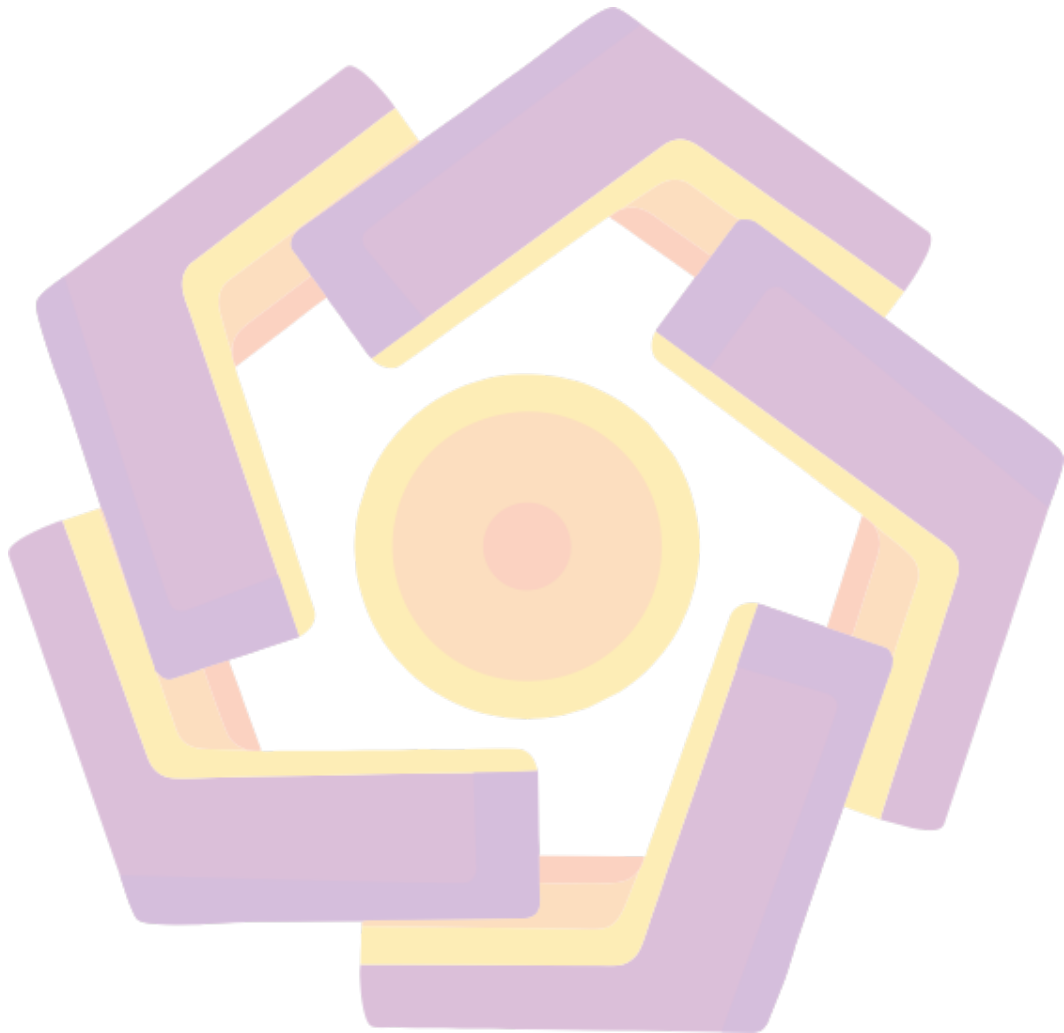
1. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Windha Mega Pradnya Duhita, M.Kom., selaku Dosen Pembimbing yang telah meluangkan waktu, tenaga, dan pikiran dalam memberikan arahan, bimbingan, serta masukan yang sangat berarti selama proses penyusunan skripsi ini.
3. Tim Dosen Penguji yang telah memberikan saran dan evaluasi demi penyempurnaan skripsi ini.
4. Seluruh Bapak/Ibu Dosen Program Studi S1 Informatika Universitas AMIKOM Yogyakarta yang telah memberikan ilmu dan pengetahuan selama masa perkuliahan.
5. Orang tua penulis yang senantiasa memberikan doa, dukungan moral, serta dukungan finansial sehingga penulis dapat menyelesaikan pendidikan dengan baik.
6. Semua pihak yang secara langsung terlibat dan berkontribusi dalam penyelesaian skripsi ini.

Penulis menyadari bahwa skripsi ini masih memiliki keterbatasan. Oleh karena itu, penulis mengharapkan saran dan masukan yang membangun demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat bagi

pembaca dan menjadi referensi yang berguna dalam pengembangan ilmu pengetahuan, khususnya di bidang *machine learning*.

Yogyakarta, 24 Februari 2026

Penulis

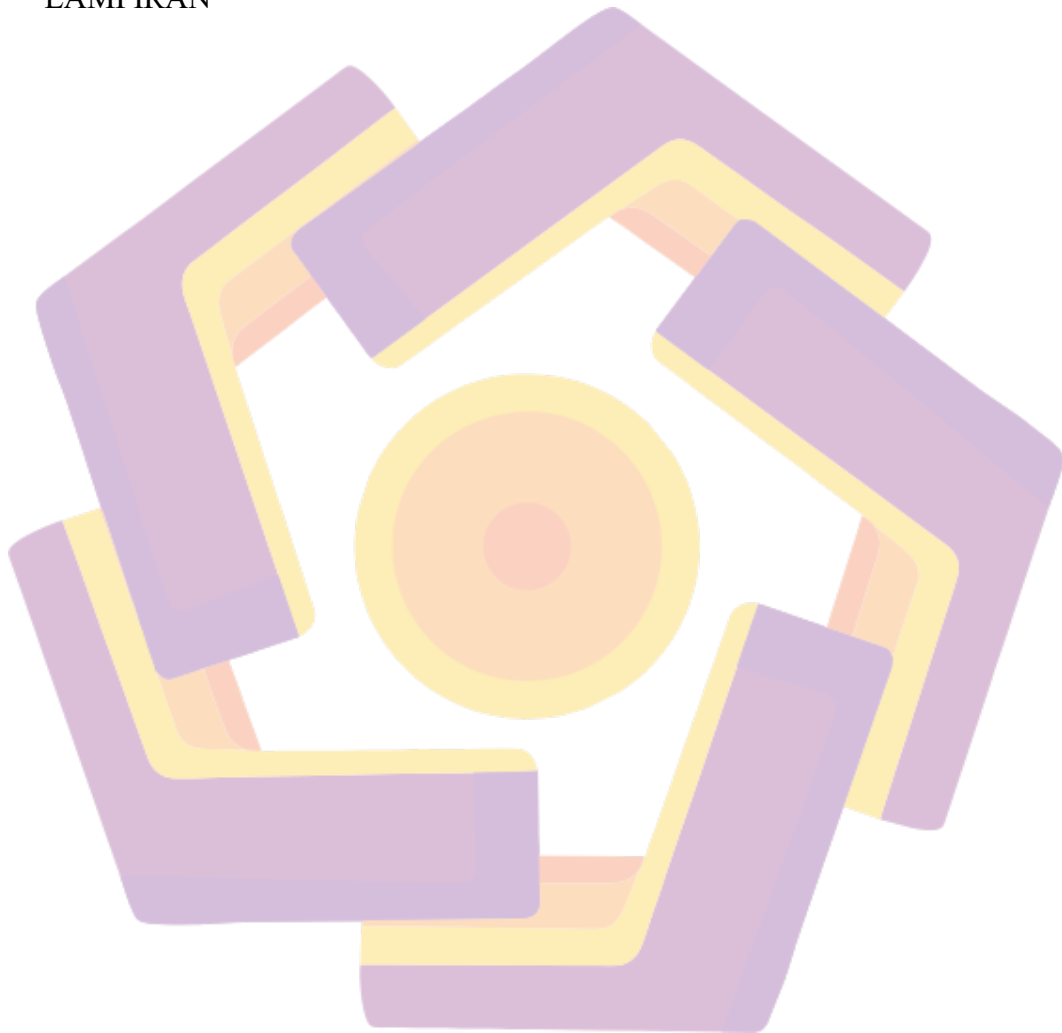


## DAFTAR ISI

|                                     |      |
|-------------------------------------|------|
| HALAMAN JUDUL                       | i    |
| HALAMAN PERSETUJUAN                 | i    |
| HALAMAN PENGESAHAN                  | ii   |
| HALAMAN PERNYATAAN KEASLIAN SKRIPSI | iii  |
| HALAMAN PERSEMBAHAN                 | iv   |
| KATA PENGANTAR                      | v    |
| DAFTAR ISI                          | vii  |
| DAFTAR TABEL                        | x    |
| DAFTAR GAMBAR                       | xi   |
| DAFTAR LAMBANG DAN SINGKATAN        | xii  |
| DAFTAR ISTILAH                      | xiii |
| INTISARI                            | xv   |
| <i>ABSTRACT</i>                     | xvi  |
| BAB I PENDAHULUAN                   | 1    |
| 1.1 Latar Belakang                  | 1    |
| 1.2 Rumusan Masalah                 | 2    |
| 1.3 Batasan Masalah                 | 2    |
| 1.4 Tujuan Penelitian               | 2    |
| 1.5 Manfaat Penelitian              | 2    |
| 1.6 Sistematika Penulisan           | 3    |
| BAB II TINJAUAN PUSTAKA             | 4    |
| 2.1 Studi Literatur                 | 4    |
| 2.2 Dasar Teori                     | 11   |
| BAB III METODE PENELITIAN           | 16   |
| 3.1. Objek Penelitian               | 16   |
| 3.2. Alur Penelitian                | 16   |
| 3.2.1. Pra penelitian               | 17   |
| 3.2.2. Pengumpulan data             | 17   |
| 3.2.3. Pra pengolahan data          | 17   |

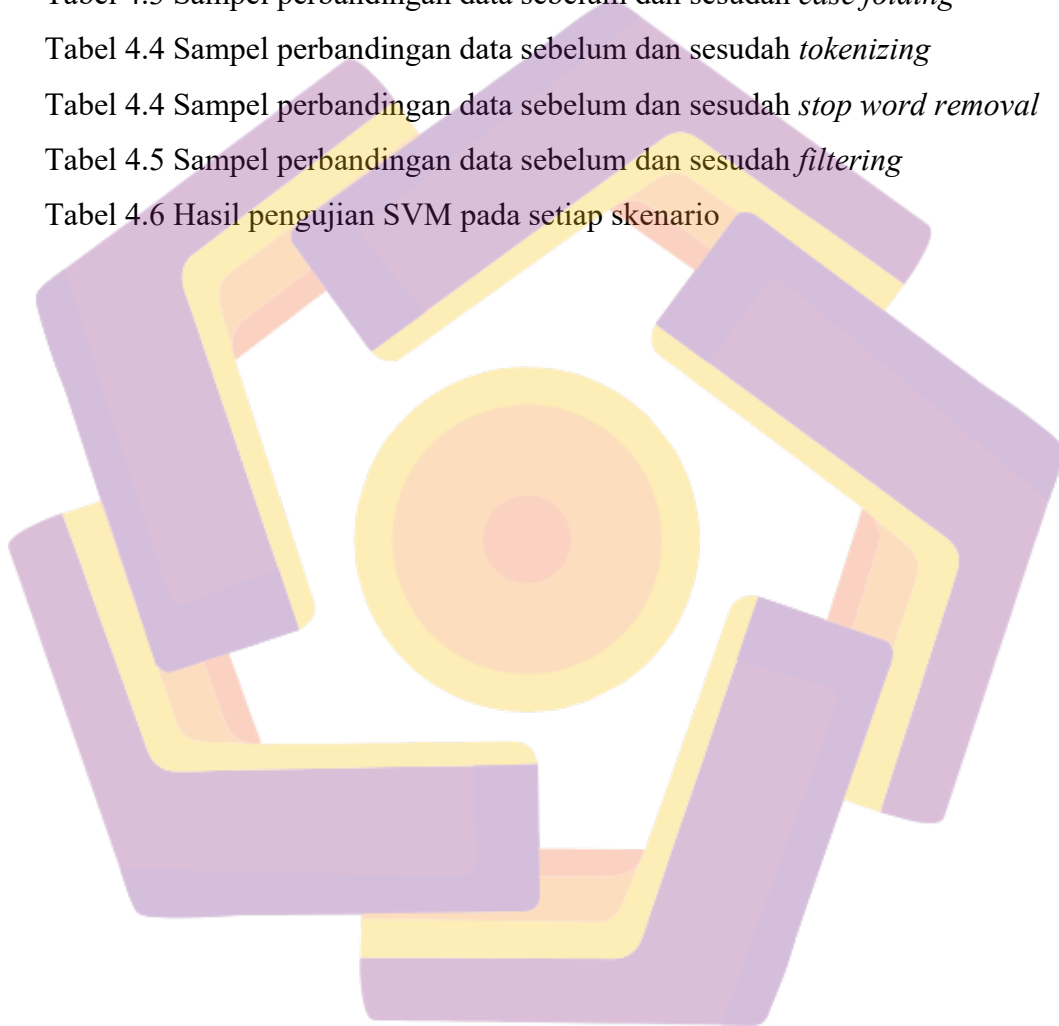
|                             |                             |    |
|-----------------------------|-----------------------------|----|
| 3.2.4.                      | Pelabelan data              | 18 |
| 3.2.5.                      | Pembobotan data             | 18 |
| 3.2.6.                      | SMOTE                       | 19 |
| 3.2.7.                      | Klasifikasi data            | 19 |
| 3.2.8.                      | Analisis atau evaluasi data | 19 |
| 3.2.9.                      | Implementasi model          | 20 |
| 3.2.10.                     | <i>Export model</i>         | 20 |
| 3.2.11.                     | <i>Build API Server</i>     | 20 |
| 3.2.12.                     | <i>Build UI Application</i> | 20 |
| 3.2.13.                     | <i>Deployment</i>           | 20 |
| 3.3.                        | Alat dan Bahan              | 21 |
| 3.3.1.                      | Perangkat lunak             | 21 |
| 3.3.2.                      | Perangkat keras             | 21 |
| BAB IV HASIL DAN PEMBAHASAN |                             | 22 |
| 4.1.                        | Pengumpulan data            | 22 |
| 4.2.                        | Pra pengolahan data         | 23 |
| 4.2.1.                      | <i>Cleansing</i>            | 24 |
| 4.2.2.                      | Normalisasi karakter        | 25 |
| 4.2.3.                      | <i>Case folding</i>         | 26 |
| 4.2.4.                      | <i>Tokenizing</i>           | 27 |
| 4.2.5.                      | <i>Stop word removal</i>    | 28 |
| 4.2.6.                      | <i>Filtering</i>            | 29 |
| 4.3.                        | Pelabelan data              | 30 |
| 4.4.                        | Pembobotan data             | 31 |
| 4.5.                        | Klasifikasi data            | 32 |
| 4.5.1.                      | Skenario tanpa SMOTE        | 33 |
| 4.5.2.                      | Skenario dengan SMOTE       | 35 |
| 4.6.                        | Analisis atau evaluasi data | 36 |
| 4.7.                        | <i>Export model</i>         | 37 |
| 4.8.                        | <i>Build API Server</i>     | 38 |
| 4.9.                        | <i>Build UI Application</i> | 39 |

|                         |    |
|-------------------------|----|
| 4.10. <i>Deployment</i> | 40 |
| BAB V PENUTUP           | 42 |
| 5.1 Kesimpulan          | 42 |
| 5.2 Saran               | 42 |
| REFERENSI               | 44 |
| LAMPIRAN                | 46 |



## DAFTAR TABEL

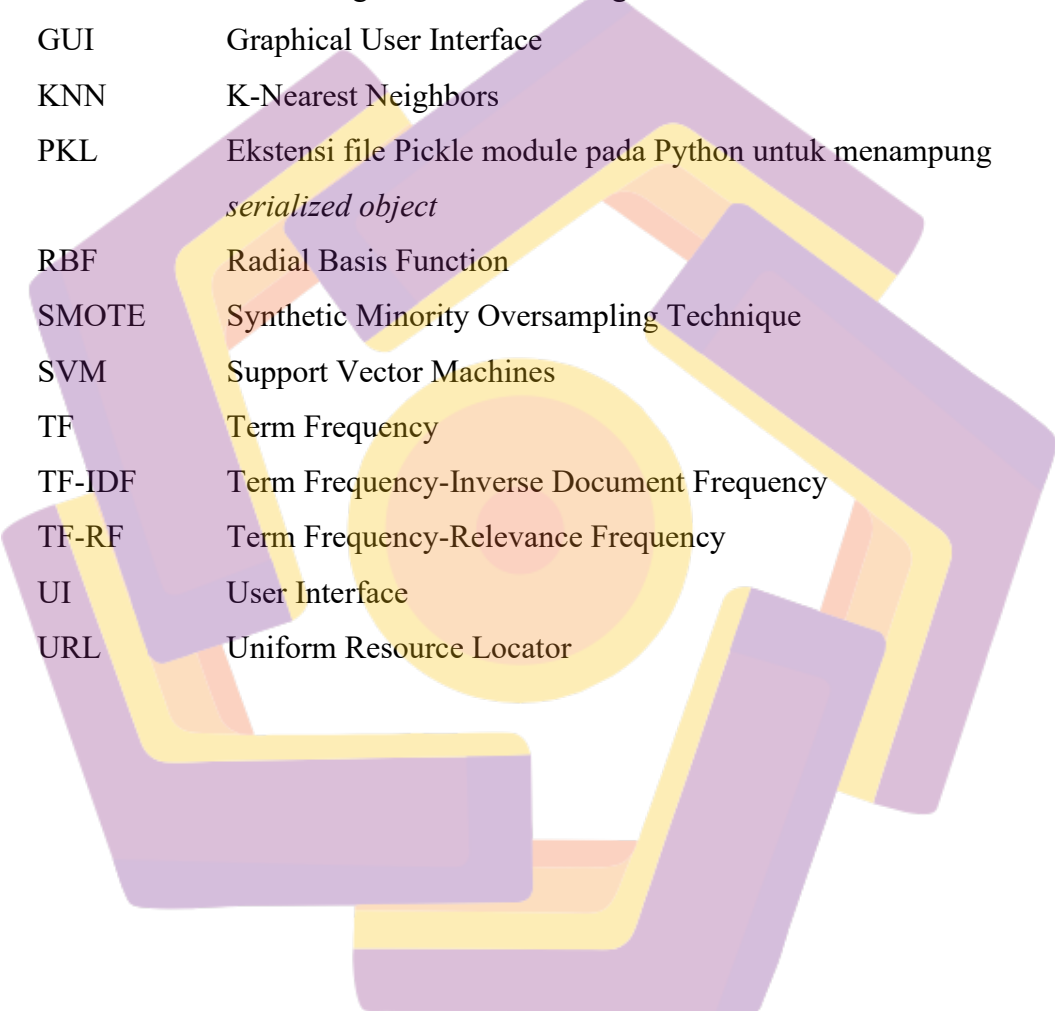
|                                                                                 |    |
|---------------------------------------------------------------------------------|----|
| Tabel 2.1 Keaslian Penelitian                                                   | 6  |
| Tabel 4.1 Sampel perbandingan data sebelum dan sesudah <i>cleansing</i>         | 24 |
| Tabel 4.2 Sampel perbandingan data sebelum dan sesudah normalisasi karakter     | 25 |
| Tabel 4.3 Sampel perbandingan data sebelum dan sesudah <i>case folding</i>      | 26 |
| Tabel 4.4 Sampel perbandingan data sebelum dan sesudah <i>tokenizing</i>        | 27 |
| Tabel 4.4 Sampel perbandingan data sebelum dan sesudah <i>stop word removal</i> | 28 |
| Tabel 4.5 Sampel perbandingan data sebelum dan sesudah <i>filtering</i>         | 29 |
| Tabel 4.6 Hasil pengujian SVM pada setiap skenario                              | 37 |



## DAFTAR GAMBAR

|                                                                                        |    |
|----------------------------------------------------------------------------------------|----|
| Gambar 3.1 Alur Penelitian                                                             | 16 |
| Gambar 4.1 Sampel <i>dataset</i> mentah                                                | 22 |
| Gambar 4.2 Plot distribusi data waktu ulasan                                           | 23 |
| Gambar 4.3 Kode program <i>method cleansing</i>                                        | 24 |
| Gambar 4.4 Kode program <i>method</i> normalisasi karakter                             | 25 |
| Gambar 4.5 Kode program <i>method case folding</i>                                     | 26 |
| Gambar 4.6 Kode program <i>method tokenizing</i>                                       | 27 |
| Gambar 4.6 Kode program <i>method stop word removal</i>                                | 28 |
| Gambar 4.7 Kode program <i>method filtering</i>                                        | 29 |
| Gambar 4.8 Kode program <i>method</i> pelabelan data                                   | 30 |
| Gambar 4.9 Sampel <i>dataset</i> pasca pengolahan                                      | 30 |
| Gambar 4.10 Plot perbandingan jenis label data                                         | 31 |
| Gambar 4.11 Sampel fitur kata pasca pembobotan data                                    | 32 |
| Gambar 4.12 Kode program skenario SVM dengan <i>kernel</i> linear tanpa SMOTE          | 33 |
| Gambar 4.13 Kode program skenario SVM dengan <i>kernel</i> RBF tanpa SMOTE             | 34 |
| Gambar 4.14 Kode program skenario SVM dengan <i>kernel</i> linear dan menerapkan SMOTE | 35 |
| Gambar 4.15 Kode program skenario SVM dengan <i>kernel</i> RBF dan menerapkan SMOTE    | 36 |
| Gambar 4.16 Kode program <i>export</i> model pra pemrosesan dan klasifikasi            | 38 |
| Gambar 4.17 Kode program untuk layanan API                                             | 38 |
| Gambar 4.18 API <i>Specification</i> yang bisa digunakan                               | 39 |
| Gambar 4.19 Kode program untuk Aplikasi <i>frontend</i>                                | 40 |
| Gambar 4.20 Sampel aplikasi pada sentimen negatif                                      | 41 |
| Gambar 4.21 Sampel aplikasi pada sentimen positif                                      | 41 |

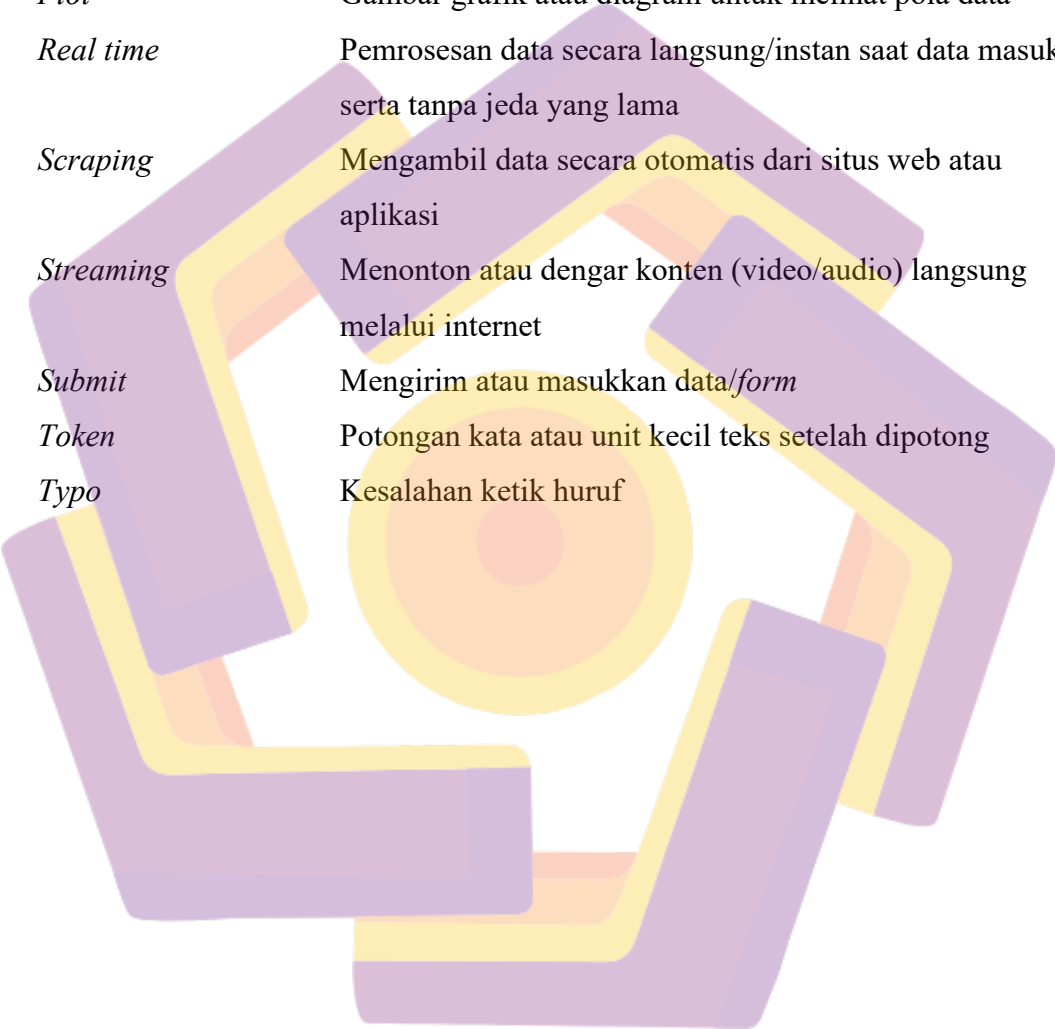
## DAFTAR LAMBANG DAN SINGKATAN



|        |                                                                                  |
|--------|----------------------------------------------------------------------------------|
| ADASYN | Adaptive Synthetic Sampling                                                      |
| API    | Application Programing Interface                                                 |
| AUC    | Area Under the Curve                                                             |
| CORS   | Cross-Origin Resource Sharing                                                    |
| GUI    | Graphical User Interface                                                         |
| KNN    | K-Nearest Neighbors                                                              |
| PKL    | Ekstensi file Pickle module pada Python untuk menampung <i>serialized object</i> |
| RBF    | Radial Basis Function                                                            |
| SMOTE  | Synthetic Minority Oversampling Technique                                        |
| SVM    | Support Vector Machines                                                          |
| TF     | Term Frequency                                                                   |
| TF-IDF | Term Frequency-Inverse Document Frequency                                        |
| TF-RF  | Term Frequency-Relevance Frequency                                               |
| UI     | User Interface                                                                   |
| URL    | Uniform Resource Locator                                                         |

## DAFTAR ISTILAH

|                     |                                                                                                                      |
|---------------------|----------------------------------------------------------------------------------------------------------------------|
| <i>Algoritma</i>    | Langkah sistematis yang diikuti komputer untuk menyelesaikan tugas                                                   |
| <i>Backend</i>      | Bagian aplikasi yang bekerja di balik layar (server)                                                                 |
| <i>Baseline</i>     | Standar awal atau patokan untuk membandingkan hasil lain                                                             |
| <i>Class</i>        | Kelompok atau kategori data                                                                                          |
| <i>Deployment</i>   | Proses memasang aplikasi/model ke server agar bisa digunakan secara umum                                             |
| <i>Emoji</i>        | Gambar yang digunakan di teks untuk menyampaikan perasaan                                                            |
| <i>Endpoint</i>     | Alamat URL spesifik di server untuk akses fitur tertentu                                                             |
| <i>Export</i>       | Menyimpan atau mengeluarkan data/model ke file lain agar bisa dipakai di luar program                                |
| <i>Field</i>        | Kolom atau bagian tertentu di tabel data                                                                             |
| <i>Fold</i>         | Bagian pembagian data saat uji coba berulang (iterasi)                                                               |
| <i>Frontend</i>     | Bagian aplikasi yang dapat dilihat pengguna                                                                          |
| <i>Hibridisasi</i>  | Menggabungkan dua metode atau algoritma berbeda                                                                      |
| <i>Hyperplane</i>   | Garis imajiner yang memisahkan kelompok data berbeda                                                                 |
| <i>Indeks</i>       | Nomor urut atau petunjuk lokasi data                                                                                 |
| <i>Kernel</i>       | Metode untuk mengubah data agar lebih mudah dipisah oleh algoritma                                                   |
| <i>Library</i>      | Kumpulan fungsi siap pakai di bahasa pemrograman                                                                     |
| <i>Method</i>       | Fungsi atau cara spesifik dalam program komputer untuk lakukan tugas tertentu                                        |
| <i>Middleware</i>   | Penghubung antara <i>frontend</i> dan <i>backend</i> , mengatur lalu lintas data, keamanan, dan koneksi antar sistem |
| <i>Noise</i>        | Data tidak terpakai yang dapat mengganggu analisis                                                                   |
| <i>Oversampling</i> | Menambah data kelompok kecil secara buatan agar seimbang dengan kelompok besar                                       |



|                  |                                                                                   |
|------------------|-----------------------------------------------------------------------------------|
| <i>Package</i>   | Bundel yang bisa diunduh untuk pemrograman                                        |
| <i>Parameter</i> | Pengaturan atau nilai yang bisa diubah untuk penyesuaian tertentu                 |
| <i>Pipeline</i>  | Rantai langkah berurutan dari awal hingga akhir proses                            |
| <i>Pip</i>       | Alat instalasi untuk unduh <i>package</i> di Python                               |
| <i>Plot</i>      | Gambar grafik atau diagram untuk melihat pola data                                |
| <i>Real time</i> | Pemrosesan data secara langsung/instan saat data masuk serta tanpa jeda yang lama |
| <i>Scraping</i>  | Mengambil data secara otomatis dari situs web atau aplikasi                       |
| <i>Streaming</i> | Menonton atau dengar konten (video/audio) langsung melalui internet               |
| <i>Submit</i>    | Mengirim atau masukkan data/form                                                  |
| <i>Token</i>     | Potongan kata atau unit kecil teks setelah dipotong                               |
| <i>Typo</i>      | Kesalahan ketik huruf                                                             |

## INTISARI

Layanan *streaming* seperti Disney Hotstar populer di Indonesia, namun fluktuasi kepuasan pengguna terlihat dari ulasan di Google Play Store yang mengandung sentimen positif dan negatif, sering kali tidak seimbang. Masalah ketidakseimbangan data ini dapat menyebabkan bias model klasifikasi, sehingga diperlukan analisis sentimen untuk mengidentifikasi kecenderungan opini pengguna guna perbaikan aplikasi. Penelitian ini bertujuan mengoptimalkan *pipeline* analisis sentimen untuk mengklasifikasikan ulasan positif-negatif secara akurat.

Metode penelitian meliputi pengumpulan data ulasan Disney Hotstar, pra-pengolahan (*cleansing*, normalisasi, *case folding*, *tokenizing*, *filtering*), pelabelan berdasarkan *rating* (positif: 4-5, negatif: 1-2), pembobotan TF-IDF, penerapan SMOTE untuk *oversampling* kelas minoritas, dan klasifikasi SVM dengan *kernel* linear serta RBF menggunakan K-Fold Cross Validation. Implementasi dilakukan dengan Python dan scikit-learn, membandingkan skenario dengan dan tanpa SMOTE serta *feature selection*. Alur ini mengoptimalkan performa model pada *dataset imbalance*.

Hasil menunjukkan akurasi model SVM mencapai sekitar 83% dengan F1-score macro 0.82-0.83, stabil meski data tidak seimbang (positif ~700, negatif ~1100). SMOTE tidak signifikan meningkatkan performa, dan *kernel* RBF sedikit lebih baik daripada linear. Model ini bermanfaat bagi developer Disney Hotstar untuk analisis ulasan dan akademisi sebagai referensi optimasi SVM pada teks Indonesia; saran lanjutan termasuk uji algoritma lain seperti XGBoost.

**Kata kunci:** analisis sentimen, SVM, SMOTE, TF-IDF, Disney.

## **ABSTRACT**

*Streaming services like Disney Hotstar are popular in Indonesia, but user satisfaction fluctuates as seen in Google Play Store reviews containing positive and negative sentiments, often imbalanced. This data imbalance can bias classification models, necessitating sentiment analysis to identify user opinion trends for app improvements. This study aims to optimize the sentiment analysis pipeline for accurately classifying positive-negative reviews.*

*The research method includes collecting Disney Hotstar reviews, preprocessing (cleansing, normalization, case folding, tokenizing, filtering), labeling based on ratings (positive: 4-5, negative: 1-2), TF-IDF weighting, SMOTE for minority class oversampling, and SVM classification with linear and RBF kernels using K-Fold Cross Validation. Implementation uses Python and scikit-learn, comparing scenarios with and without SMOTE and feature selection. This flow optimizes model performance on imbalanced datasets.*

*Results show SVM model accuracy around 83% with macro F1-score 0.82-0.83, stable despite imbalance (positive ~700, negative ~1100). SMOTE provided no significant improvement, and RBF kernel slightly outperformed linear. The model benefits Disney Hotstar developers for sentiment analysis and academics as SVM optimization reference for Indonesian text; future suggestions include testing other algorithms like XGBoost.*

**Keyword:** *sentiment analysis, SVM, SMOTE, TF-IDF, Disney Hotstar*