

**ANALISIS SENTIMEN FILM DI TWITTER
MENGUNAKAN METODE SUPPORT
VECTOR MACHINE (SVM)**

TUGAS AKHIR

Disusun sebagai salah satu syarat untuk kelulusan Program Strata 1,
di Program Studi Teknik Informatika, Universitas Pasundan Bandung

oleh :

Mahdy Khairudin
NPM : 17.304.0146



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS PASUNDAN BANDUNG
SEPTEMBER 2021**

LEMBAR PENGESAHAN
LAPORAN TUGAS AKHIR

Telah diujikan dan dipertahankan dalam Sidang Sarjana Program Studi Teknik Informatika Universitas Pasundan Bandung, pada hari dan tanggal sidang sesuai berita acara sidang, tugas akhir dari :

Nama : Mahdy Khairudin

Nrp : 17.304.0146

Dengan judul :

**“ANALISIS SENTIMEN FILM DI TWITTER MENGGUNAKAN METODE
SUPPORT VECTOR MACHINE (SVM)”**



Bandung, 23 September 2021

Menyetujui,

Pembimbing Utama,

(Ade Sukendar, ST., MT.)

ABSTRAK

Twitter adalah salah satu media sosial yang sering digunakan untuk berbagi informasi atau biasa disebut *tweet* secara *realtime*. Berdasarkan hasil penelitian pada tahun 2013 oleh SemioCast, lembaga riset media sosial, mengatakan bahwa jumlah pengguna *Twitter* di Indonesia terbanyak ketiga setelah Amerika Serikat dan Jepang dengan angka 6,5 persen pengguna dan 1 miliar *tweet*.

Tweet biasanya dapat mempengaruhi minat beli, memilih suatu tokoh politik, atau menonton film mereka. *Tweet* berkaitan dengan film menjadi referensi bagi pengguna *Twitter* yang lain apabila ingin menonton film yang sama. Karena tidak adanya layanan yang menyediakan referensi film berdasarkan komentar orang lain yang relevan untuk penonton Indonesia.

Penelitian ini untuk melakukan analisis sentimen terhadap *tweet* di *Twitter* pada kasus film berbahasa Indonesia yang menghasilkan *rating* film. Metode yang digunakan untuk analisis sentimen adalah dengan *Support Vector Machine* (SVM)

Hasil akhir dari penelitian ini adalah sebuah model atau algoritma SVM yang dapat mengklasifikasikan atau analisis sentimen film berdasarkan *tweet* di *Twitter* dengan memanfaatkan *Twitter API*.

Kata kunci: analisis sentimen, *Support Vector Machine*, *Twitter API*, dan film

ABSTRACT

Twitter is one of the social media that is often used to share information or so-called tweets in real time. Based on the results of research in 2013 by SemioCast, a social media research institute, said that the number of Twitter users in Indonesia was third after the United States and Japan with 6.5 percent of users and 1 billion tweets.

Tweets can usually influence buying interest, voting for a political figure, or watching their movie. Tweets related to movies become a reference for other Twitter users if they want to watch the same movie. Because there is not service that provides movie references based on other people's comments that are relevant for Indonesian audiences.

This study is to analyze the sentiment of tweets on Twitter in the case of Indonesian-language movies that produce movie ratings. The method used for sentiment analysis is the Support Vector Machine (SVM).

The final result of this research is a model or SVM algorithm that can classify or analyze film sentiment based on tweets on Twitter by utilizing the Twitter API.

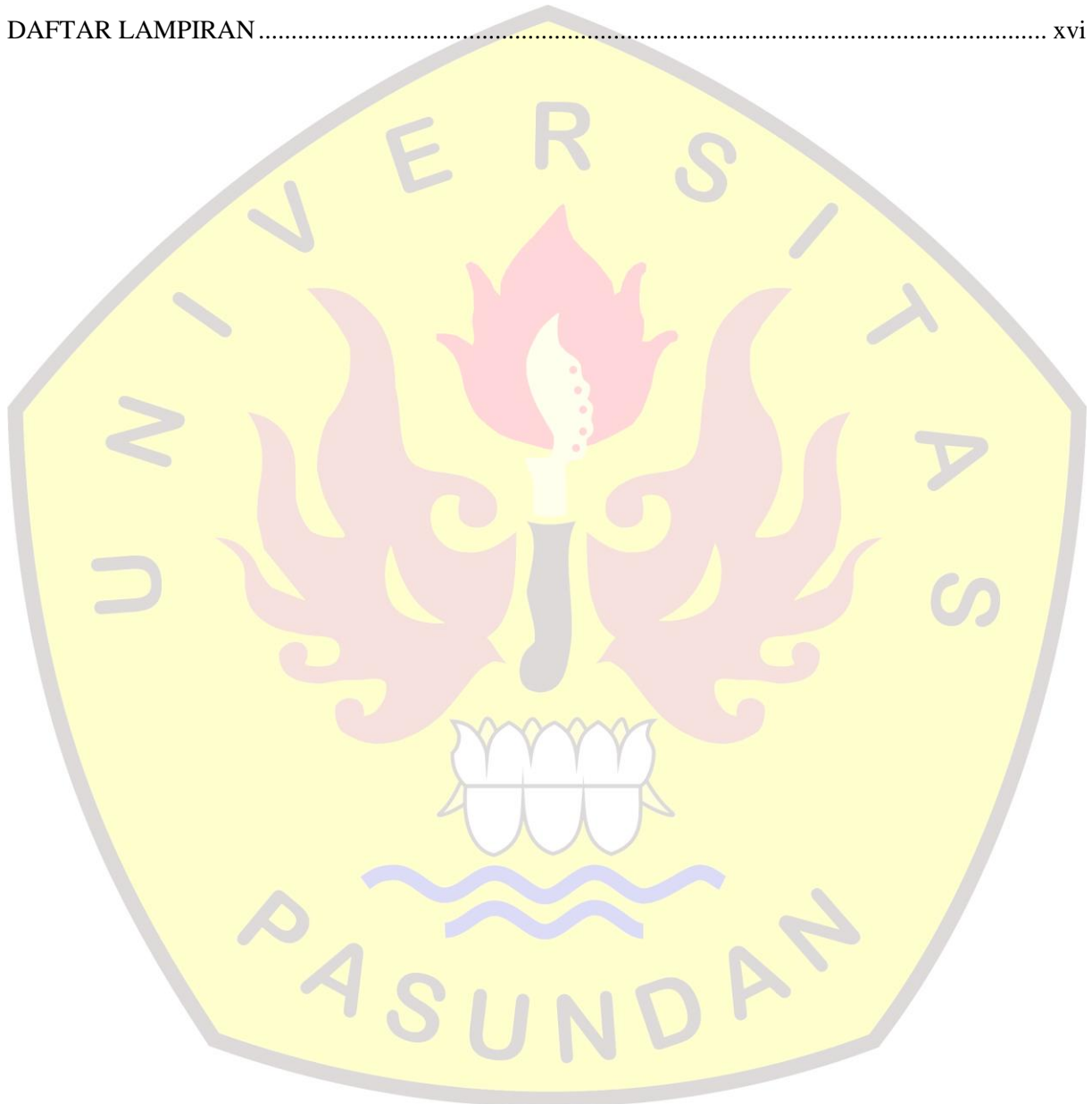
Keywords: sentiment analysis, Support Vector Machine, Twitter API, and movies

DAFTAR ISI

LEMBAR PENGESAHAN.....	ii
LAPORAN TUGAS AKHIR.....	ii
LEMBAR PERNYATAAN KEASLIAN TUGAS AKHIR	iii
ABSTRAK.....	iv
ABSTRACT.....	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR	xi
DAFTAR ISTILAH	xiii
DAFTAR SIMBOL.....	xiv
BAB 1 PENDAHULUAN	1-1
1.1 Latar Belakang	1-1
1.2 Identifikasi Masalah	1-2
1.3 Tujuan Tugas Akhir	1-2
1.4 Lingkup Tugas Akhir	1-2
1.5 Metodologi Tugas Akhir	1-2
BAB 2 LANDASAN TEORI DAN PENELITIAN TERDAHULU.....	2-1
2.1 Landasan Teori.....	2-1
2.1.1 <i>Data Mining</i>	2-1
2.1.2 <i>Twitter API</i>	2-2
2.1.3 Analisis Sentimen.....	2-3
2.1.4 <i>Preprocessing</i>	2-4
2.1.5 Algoritma Klasifikasi	2-5
2.1.6 Algoritma <i>Support Vector Machine</i>	2-7
2.2 Penelitian Terdahulu	2-9
BAB 3 SKEMA PENELITIAN	3-1

3.1 Alur Penyelesaian Tugas Akhir.....	3-1
3.2 Perumusan Masalah.....	3-4
3.2.1 Analisis Sebab Akibat	3-4
3.2.2 Solusi Masalah	3-5
3.3. Kerangka Pemikiran Teoritis	3-5
3.3.1 Gambaran Produk TA	3-5
3.3.2 Skema Analisis Teori	3-7
3.4 Profile Penelitian.....	3-10
3.4.1 Objek Penelitian	3-10
BAB 4 STUDI ANALISIS SENTIMEN FILM DI TWITTER	4-1
4.1 Pengumpulan Data	4-1
4.2 Preprocessing data.....	4-2
4.3 Melakukan Analisis Sentimen.....	4-7
4.3.1 Pembuatan Model SVM.....	4-7
4.3.2 Pengklasifikasian Sentimen Data	4-11
4.3.3 Penghitungan Rating	4-12
BAB 5 IMPLEMENTASI PROTOTYPE ANALISIS SENTIMEN FILM.....	5-1
5.1 Analisis dan Desain.....	5-1
5.1.1 <i>User Requirements</i>	5-1
5.1.2 <i>Software Requirements</i>	5-1
5.1.3 Model Use Case	5-2
5.1.4 Skenario <i>Use Case</i>	5-3
5.1.4 <i>Sequence Diagram</i>	5-4
5.1.5 <i>Component Diagram</i>	5-5
5.1.6 Struktur Data	5-6
5.2 Implementasi.....	5-8
5.1.1 Kakas yang digunakan (<i>Tools</i>).....	5-8
5.1.2 Struktur Project Aplikasi.....	5-9
5.1.3 Prototype Aplikasi Web	5-11

5.3 Pengujian.....	5-13
BAB 6 KESIMPULAN DAN SARAN.....	6-1
6.1 Kesimpulan	6-1
6.2 Saran.....	6-1
DAFTAR PUSTAKA	xv
DAFTAR LAMPIRAN.....	xvi



BAB 1

PENDAHULUAN

1.1 Latar Belakang

Twitter adalah salah satu media sosial yang sering digunakan untuk berbagi informasi atau biasa disebut *tweet* secara *realtime*. Berdasarkan hasil penelitian pada tahun 2013 oleh SemioCast, lembaga riset media sosial, mengatakan bahwa jumlah pengguna *Twitter* di Indonesia terbanyak ketiga setelah Amerika Serikat dan Jepang dengan angka 6,5 persen pengguna dan 1 miliar *tweet* [2].

Pengguna *Twitter* yang semakin meningkat menimbulkan peningkatan *tweet* yang diposting. *Twitter* juga adalah salah satu media sosial yang efektif dalam menampung opini dari mengenai produk, layanan termasuk film karena dalam memberikan komentar *Twitter* sangat cepat dan hanya berupa teks mirip seperti melakukan SMS. Dengan *Twitter* memudahkan dalam mengumpulkan data opini dibandingkan dengan survey maupun penyebaran kuisioner.

Data mining adalah analisis kumpulan data observasi (seringkali memiliki volume data besar) untuk menemukan hubungan yang tidak terduga dan untuk meringkas data dengan cara baru yang dapat dimengerti dan berguna bagi pemilik data [7]. Dari definisi tersebut *data mining* berhubungan dengan pengelolaan data yang jumlah datanya besar dan datanya tidak terstruktur.

Analisis sentimen atau disebut juga dengan *opinion mining* adalah salah satu bidang yang melakukan studi mengenai opini orang-orang, sentimen, evaluasi, tingkah laku dan emosi dari seseorang terhadap suatu entitas seperti produk, layanan, organisasi, individu, permasalahan, topik, acara dan atribut-atributnya [4]. Analisis sentimen berhubungan dengan pengelolaan data yang berkaitan dengan opini orang lain terhadap suatu entitas. Dengan kata lain analisis sentimen adalah sub bidang dari *data mining*.

Pada studi kasus film, *tweet* biasanya dapat mempengaruhi minat orang dalam menonton film tertentu. *Tweet* berkaitan dengan film menjadi referensi bagi pengguna *Twitter* yang lain apabila ingin menonton film yang sama. Karena tidak adanya layanan yang menyediakan referensi film berdasarkan komentar orang lain yang relevan untuk penonton Indonesia.

Celah atau Masalah inilah yang kemudian penulis buat untuk mengimplementasikan analisis sentimen pada *Twitter* sebagai pemberi *rating* film.

1.2 Identifikasi Masalah

Berdasarkan pemaparan dari latar belakang diatas, maka permasalahan yang dimunculkan pada tugas akhir ini adalah sebagai berikut:

1. Bagaimana mengumpulkan data dari *tweet* berbahasa Indonesia di *Twitter*?
2. Bagaimana membuat model SVM yang dapat menganalisis sentimen dari *tweet*?
3. Bagaimana menghitung *rating* film dari hasil model SVM?

1.3 Tujuan Tugas Akhir

Tujuan yang ingin dicapai dari tugas akhir ini sebagai berikut:

1. Untuk membuat model dan *prototype* analisis sentimen menggunakan algoritma SVM terhadap data komentar di *Twitter* pada kasus film berbahasa Indonesia yang menghasilkan *rating* film.

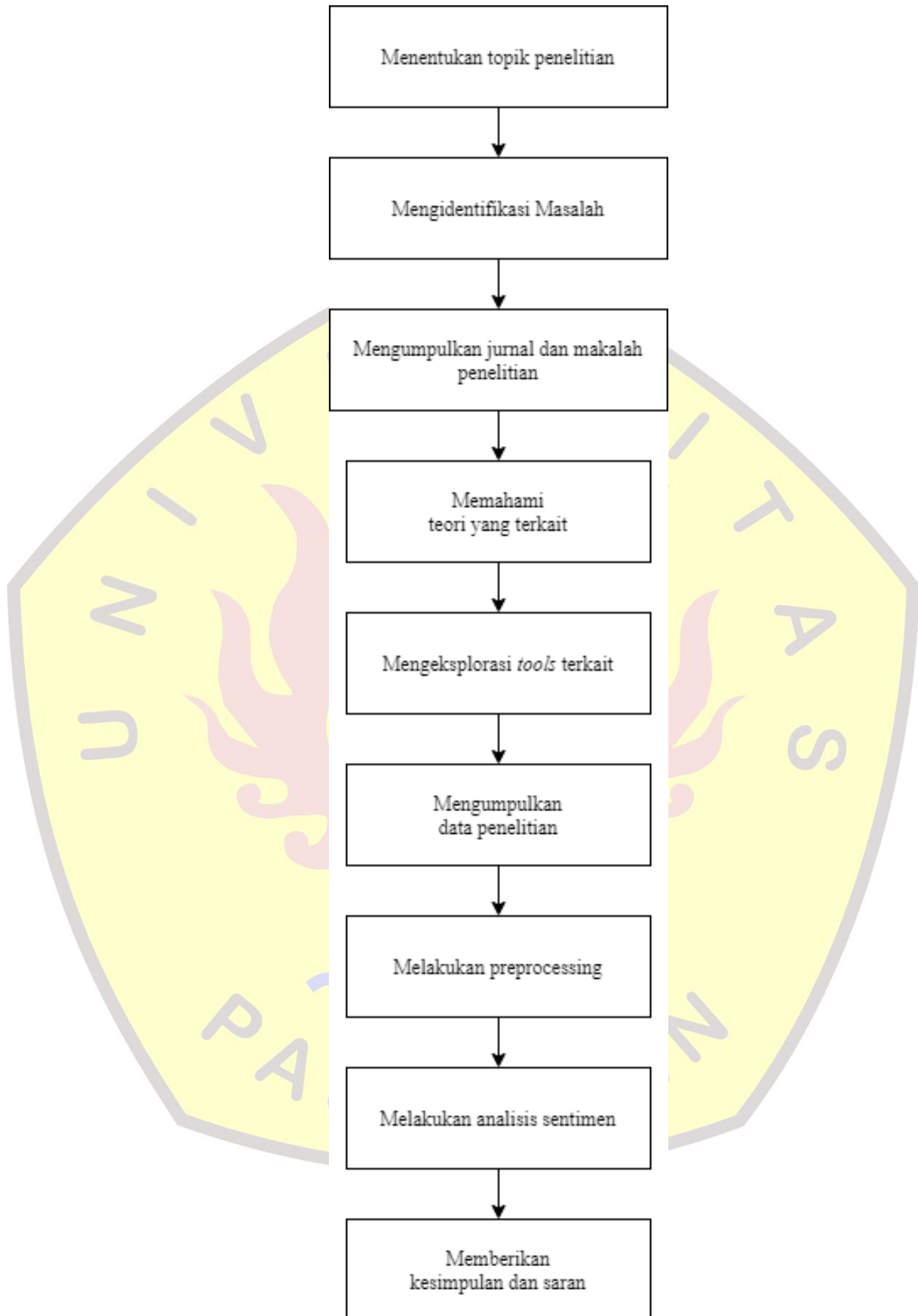
1.4 Lingkup Tugas Akhir

Penyelesaian tugas akhir dibatasi sebagai berikut:

1. Pengumpulan data dari komentar atau *tweet* yang digunakan adalah berbahasa Indonesia.
2. Pengumpulan data untuk studi kasusnya menggunakan media sosial *Twitter*.
3. Tahapan yang dilakukan pada penelitian ini adalah dari *crawling data* sampai penghitungan *rating* untuk film.
4. Studi kasus yang digunakan adalah film.
5. Analisis sentimen untuk mengklasifikasikan komentar atau *tweet* menggunakan model *Support Vector Machine* (SVM).
6. *Data training* dan *data testing* yang digunakan untuk membuat model dari penelitian berjudul “Analisis Sentimen Tentang Opini Film Pada Dokumen *Twitter* Berbahasa Indonesia Menggunakan *Naive Bayes* Dengan Perbaikan Kata Tidak Baku” dibuat oleh Rizal Setya Perdana [12]. Data tersebut berjumlah 200 data.
7. Pembagian *data training* dan *data testing* sebesar 80% *data training* dan 20% *data testing*.
8. *Twitter API* yang digunakan API yang standar atau gratis sehingga pencarian *query* terbatas hanya seminggu ke belakang.
9. Sentimen yang diklasifikasikan yaitu positif dan negatif.

1.5 Metodologi Tugas Akhir

Bagian ini menjelaskan langkah-langkah metodologi penyelesaian yang dilakukan penulis dalam menyelesaikan tugas akhir. Langkah-langkah tersebut dijelaskan dalam bentuk diagram yang dapat dilihat pada gambar 1.1.



Gambar 1. 1 Metodologi Penyelesaian Tugas Akhir

Berikut merupakan rincian dari metodologi penyelesaian tugas akhir ini, diantaranya:

1. Menentukan topik penelitian

Pada tahap ini dilakukan untuk menentukan topik penelitian yang akan diajukan oleh penulis.

2. Mengidentifikasi masalah

Pada tahap ini dilakukan untuk melakukan identifikasi masalah yang terjadi terkait dengan topik penelitian, serta solusi sementara yang akan diusulkan untuk mengatasi masalah tersebut.

3. Mengumpulkan jurnal dan makalah penelitian

Pada tahap ini dilakukan untuk mengumpulkan referensi yang relevan terkait masalah dan solusi sementara yang akan diusulkan.

4. Memahami teori yang terkait

Pada tahap ini dilakukan untuk memahami teori dan konsep yang terkait dengan masalah yang sudah diidentifikasi dan referensi yang didapatkan.

5. Mengeksplorasi *tools* terkait

Pada tahap ini dilakukan untuk mencari dan mengeksplorasi *tools* terkait dengan topik penelitian yang akan digunakan pada penelitian tersebut.

6. Mengumpulkan data penelitian

Pada tahap ini dilakukan untuk mengumpulkan data penelitian berupa data dari *tweet* berbahasa Indonesia di *Twitter* yang berkaitan dengan film dengan menggunakan *Twitter API* dan data *training* serta *testing* untuk membuat model analisis sentimen dari data penelitian ini [12].

7. Melakukan *preprocessing*

Pada tahap ini dilakukan *preprocessing* data agar data dari *Twitter API* dan data *training* serta data *testing* yang didapatkan menjadi terstruktur, bahasanya menjadi baku dan menghilangkan kesalahan penulisan. Dengan *preprocessing* ini membuat data dapat diolah oleh algoritma *Support Vector Machine* nya.

8. Melakukan analisis sentimen

Pada tahap ini dilakukan analisis sentimen data yang terdiri dari membuat model analisis sentimen dengan algoritma *Support Vector Machine*, menganalisis sentimen data film dari *Twitter API* dengan model yang telah dibuat dan pengitungan rating dari film.

9. Memberikan kesimpulan dan saran

Pada tahap ini dilakukan untuk memberikan kesimpulan dan saran dari hasil penelitian yang telah dilakukan.

1.6 Sistematika Penulisan Tugas Akhir

Sistematika penulisan yang menjelaskan mengenai bab-bab pada laporan tugas akhir beserta isinya secara rinci, serta keterkaitan antara bab sebelum dan sesudahnya. Adapun sistematika penulisannya adalah sebagai berikut:

BAB 1. PENDAHULUAN

Bab ini berisi penjelasan umum mengenai usulan penelitian yang dilakukan dalam pengerjaan tugas akhir. Di dalamnya berisi latar belakang masalah, identifikasi masalah, tujuan tugas akhir, lingkup tugas akhir, metodologi pengerjaan tugas akhir, dan sistematika penulisan laporan tugas akhir.

BAB 2. LANDASAN TEORI DAN PENELITIAN TERDAHULU

Bab ini berisi definisi, teori-teori serta konsep yang diperlukan dalam pengerjaan tugas akhir. Bab ini juga membahas mengenai jurnal-jurnal ilmiah terdahulu yang memiliki kemiripan dengan tugas akhir yang dikerjakan.

BAB 3. SKEMA PENELITIAN

Bab ini berisi penjelasan alur penyelesaian tugas akhir, analisis persoalan dan manfaat tugas akhir, kerangka pemikiran teoritis, dan profil tempat penelitian.

BAB 4. STUDI ANALISIS SENTIMEN FILM DI TWITTER

Bab ini berisi studi analisis sentimen film di twitter dengan menggunakan salah satu algoritma klasifikasi yaitu *Support Vector Machine (SVM)*

BAB 5. IMPLEMENTASI PROTOTYPE ANALISIS SENTIMEN FILM DI TWITTER

Bab ini berisi implementasi *prototype* analisis sentimen film di twitter pada berbasis web menggunakan *framework* Django.

BAB 6. PENUTUP

Bab ini berisi mengenai hasil penelitian serta pernyataan yang didapat berdasarkan identifikasi masalah yang dikemukakan, serta keterkaitan dari semua tahap yang dilakukan dalam penelitian. Di dalamnya terdapat pula saran yang diusulkan untuk penelitian selanjutnya terkait dengan prospek penelitian selanjutnya.

DAFTAR PUSTAKA

- [1] Akshay Amolik, et al., 2016. *Twitter Sentiment Analysis of Movie Reviews using Machine Learning Techniques*, International Journal of Engineering and Technology (IJET).
- [2] Rindu P. Hesty, 2013. Indonesia Pengguna Twitter Nomor 3 di Dunia. [Online]. <https://tekno.tempo.co/read/538043/indonesia-pengguna-twitter-nomor-3-di-dunia> [Diakses 18 Agustus 2020].
- [3] Nugroho, Witarto, & Handoko, 2003. *Application of Support Vector Machine in Bioinformatics*. Proceesing of Indonesian Scientific Meeting in Central.
- [4] B. Liu, 2012. *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers
- [5] Eyheramendy, S., Lewis, D. D., & Madigan, D. 2013. *On the Naive Bayes Model for Text Categorization. Ninth International Workshop on Artificial Intellegence and Statistics*.
- [6] Hadna, N. et al., 2016. Studi Literatur Tentang Perbandingan Metode Proses Analisis Sentimen di Twitter. Fakultas Teknik, Universitas Gadjah Mada.
- [7] D. Hand, H. Mannila, & P. Smyth, 2001. *Principles of Data Mining*. A Bradford Book The MIT Press.
- [8] Buntoro, G. A., Adji, T. B., & Purnamasari, A. E. 2014. Sentiment Analysis Twitter dengan Kombinasi Lexicon Based dan Double Propagation. CTIEE.
- [9] Twitter, "Twitter API Docs", [Online]. Available: <https://developer.twitter.com/en/docs> [Diakses 20 Maret 2021].
- [10] Jiawei Han, M. Kamber, Jian Pei, 2012, "*Data Mining Concepts and Techniques Third Edition*". Morgan Kaufmann Publisher.
- [11] Nuvirta M., Lukito E. N., Bimo S. H., 2014. *Penerapan Analisis Sentimen Pada Twitter Berbahasa Indonesia Sebagai Pemberi Rating*. Jurusan Teknologi Informasi, Universitas Gajah Mada.
- [12] Prananda A., Rizal S. P., M. Ali Fauzi, 2017. *Analisis Sentimen Tentang Opini Film Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naive Bayes Dengan Perbaikan Kata Tidak Baku*. Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Brawijaya.
- [13] Ian Sommerville, 2011. *Software Engineering, 9th Edition*. Boston, Massachusetts: Pearson Education.