

**REKOMENDASI MASKAPAI PENERBANGAN BERDASARKAN OPINI  
MASYARAKAT DI TWITTER DENGAN BERBASIS SENTIMEN  
ANALISIS MENGGUNAKAN ALGORITMA SVM**

**SKRIPSI**



disusun oleh

**Heriyono**

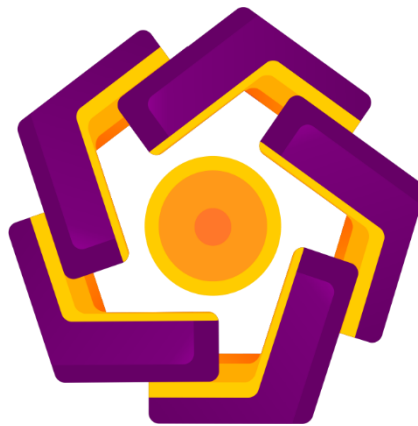
**15.11.8738**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2022**

**REKOMENDASI MASKAPAI PENERBANGAN BERDASARKAN OPINI  
MASYARAKAT DI TWITTER DENGAN BERBASIS SENTIMEN  
ANALISIS MENGGUNAKAN ALGORITMA SVM**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Informatika



disusun oleh

**Heriyono**

**15.11.8738**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2022**

**PERSETUJUAN**

**SKRIPSI**

**REKOMENDASI MASKAPAI PENERBANGAN BERDASARKAN OPINI  
MASYARAKAT DI TWITTER DENGAN BERBASIS SENTIMEN  
ANALISIS MENGGUNAKAN ALGORITMA SVM**

yang dipersiapkan dan disusun oleh

**Heriyono**

**15.11.8738**

telah disetujui oleh Dosen Pembimbing Skripsi

pada tanggal 02 Juni 2022

**Dosen Pembimbing,**

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

**NIK. 190302231**

# PENGESAHAN

## SKRIPSI

### REKOMENDASI MASKAPAI PENERBANGAN BERDASARKAN OPINI MASYARAKAT DI TWITTER DENGAN BERBASIS SENTIMEN ANALISIS MENGGUNAKAN ALGORITMA SVM

yang dipersiapkan dan disusun oleh

**Heriyono**  
**15.11.8738**

telah dipertahankan di depan Dewan Penguji

pada tanggal 23 Juni 2022

**Susunan Dewan Penguji**

**Nama Penguji**

**Dina Maulina, M.Kom**

**NIK. 190302250**

**Eli Pujastuti, M.Kom**

**NIK. 190302227**

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

**NIK. 190302231**

**Tanda Tangan**

Skripsi ini telah diterima sebagai salah satu persyaratan

untuk memperoleh gelar Sarjana Komputer

Tanggal 23 Juni 2022

**Dekan Fakultas Ilmu Komputer**

**Hanif Al Fatta, M.Kom**

**NIK. 190302096**

## 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, 23 Juni 2022



Heriyono

NIM. 15.11.8738

## MOTTO

*”Sehat tidak menarik ketika kita miliki, Sehat Menjadi menarik ketika pergi dari kita ”*

**-Ade Rai**

*”Kemajuan Kecil, Tetaplah Kemajuan”*

**-Morgan Harper Nichols**

*”Perbaiki Sholatmu maka Allah akan perbaiki hidupmu”*

**-Heri**

## PERSEMBAHAN

Puji Syukur saya panjatkan kepada Allah SWT karena telah berhasil menyelesaikan Skripsi ini. Saya persembahkan Skripsi ini untuk:

- Ayah dan Ibu Tercinta, Karena kalian berdua hidup terasa begitu mudah dan penuh dengan kebahagiaan. Terima kasih karena telah menjaga saya dalam doa-doa ayah dan ibu serta selalu membiarkan saya mengejar impian saya apapun itu.
- Adikku terima kasih untuk dukungannya selama ini.
- Dosen Pembimbing Ibu Erni Seniwati, S.Kom, M.Cs yang telah membimbing saya dalam pembuatan tugas akhir ini.
- Teman-teman kos putra muslim yang selalu memberikan dukungan.
- Teman-teman TI-04 terima kasih atas dukungannya.
- Dosen Universitas Amikom Yogyakarta yang telah memberikan ilmunya kepada saya selama menempuh kuliah.
- Diri saya sendiri, saya mengapresiasi diri saya sendiri atas pembuatan penelitian ini.

## KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah subhanahu wa ta'ala yang telah melimpahkan nikmat dan rahmat-Nya yang melimpah, sehingga penulis dapat menyelesaikan Skripsi yang berjudul “*Rekomendasi Maskapai Penerbangan Berdasarkan Opini Masyarakat Di Twitter Dengan Berbasis Sentimen Analisis Menggunakan Algoritma SVM*” dengan tepat waktu.

Penulisan Skripsi ini dimaksudkan untuk memenuhi syarat kelulusan program Sarjana Informatika di Universitas Amikom Yogyakarta.


Selesaiannya skripsi ini tidak terlepas dari dukungan berbagai pihak yang telah memberikan dorongan moril maupun spiritual dan juga bimbingan ilmu pengetahuan. Oleh karena itu penulis mengucapkan terimakasih kepada :

1. Kedua orang tua dan adik penulis sebagai sosok luar biasa yang memberikan segala dukungan dan semangat baik moril maupun material.
2. Bapak Prof. Dr. M. Suyanto, MM selaku rektor Universitas Amikom Yogyakarta.
3. Ibu Erni Seniwati, S.Kom, M.Cs selaku dosen pembimbing yang telah membimbing dan mengarahkan dalam pembuatan skripsi ini.
4. Seluruh dosen, staff maupun karyawan Universitas Amikom Yogyakarta.
5. Teman-teman di Universitas Amikom Yogyakarta yang telah memberikan dukungan.



6. Semua pihak yang telah banyak membantu dalam penyusunan skripsi ini yang tidak dapat disebutkan satu persatu.

Penulis menyadari sepenuhnya bahwa skripsi ini banyak mengandung kekurangan , untuk itu segala kritik dan saran yang membangun sangat diharapkan. Akhir Kata, penulis berharap semoga makalah ini dapat memberikan manfaat bagi pihak yang memerlukan.



Yogyakarta, 23 Juni 2022



**Heriyono**

## DAFTAR ISI

<b>PERSETUJUAN</b> .....	<b>II</b>
<b>PENGESAHAN</b> .....	<b>III</b>
<b>PERNYATAAN</b> .....	<b>IV</b>
<b>MOTTO</b> .....	<b>V</b>
<b>PERSEMBAHAN</b> .....	<b>VI</b>
<b>KATA PENGANTAR</b> .....	<b>VII</b>
<b>DAFTAR ISI</b> .....	<b>IX</b>
<b>DAFTAR TABEL</b> .....	<b>XIII</b>
<b>DAFTAR GAMBAR</b> .....	<b>XIV</b>
<b>INTISARI</b> .....	<b>XVI</b>
<b>ABSTRACT</b> .....	<b>XVII</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Penelitian.....	2
1.4 Maksud dan Tujuan Penelitian.....	3
1.5 Metodologi Penelitian.....	4
1.5.1 Metode Pengumpulan Data.....	4
1.5.2 Tahapan Penelitian.....	4
1.6 Sistematika Penulisan .....	7
<b>BAB II LANDASAN TEORI</b> .....	<b>9</b>
2.1 Tinjauan Pustaka.....	9

2.2	Pengertian Data Mining .....	11
2.3	Pengertian Analisis Sentimen .....	12
2.4	Pre-processing.....	13
2.5	Term Frequency-Inverse Document Frequency (TF-IDF.....	13
2.6	Cosine Similarity .....	14
2.7	Support Vector Machine .....	14
2.8	K-Fold Validation .....	17
2.9	Confusion Matrix .....	18
<b>BAB III ANALISIS DAN PERANCANGAN.....</b>		<b>20</b>
3.1	Analisis Masalah.....	20
3.2	Solusi Yang Dipilih.....	21
3.3	Analisis Kebutuhan.....	22
3.3.1	Analisa Kebutuhan Fungsional.....	22
3.3.2	Analisa Kebutuhan Non-Fungsional.....	22
3.4	Analisis Sentimen dengan Klasifikasi Algoritma SVM .....	23
3.4.1	Data Cleaning .....	23
3.4.2	Case Folding .....	25
3.4.3	Tokenization .....	26
3.4.4	Stopword.....	27
3.4.5	TF-IDF .....	29
3.4.6	Cosine Similarity .....	32
3.4.7	Support Vector Machine.....	38

3.4.8	K-Fold Cross Validation.....	40
3.4.9	Confusion Matrix.....	41
3.5	Perancangan Aplikasi.....	42
3.5.1	Diagram Alir Pre-processing.....	42
3.5.2	Proses Klasifikasi Support Vector Machine.....	43
3.6	Perancangan Crawling Data Twitter.....	45
3.7	Perancangan File Training.....	45
3.8	Langkah-langkah Pengembangan Aplikasi.....	46
<b>BAB IV IMPLEMENTASI DAN PEMBAHASAN.....</b>		<b>50</b>
4.1	Deskripsi Implementasi.....	50
4.2	Implementasi Pengumpulan Data Tweet.....	50
4.3	Implementasi Pelabelan Data Training.....	53
4.4	Implementasi Perancangan Script dan Sistem.....	54
4.4.1	Pembuatan Script Akses Data Training.....	54
4.4.2	Pembuatan Script Cleaning dan Case Folding.....	54
4.4.3	Pembuatan Script Tokenization.....	55
4.4.4	Pembuatan Script Stopword.....	56
4.4.5	Implementasi Pembuatan Fungsi Klasifikasi SVM.....	56
4.4.6	Implementasi Pembuatan Prediksi Sentimen.....	57
4.5	Implementasi Antarmuka.....	57
4.6	Evaluasi Program.....	59
4.6.1	Pengujian Algoritma.....	59

<b>BAB V KESIMPULAN .....</b>	<b>66</b>
5.1 Kesimpulan .....	66
5.2 Saran .....	66
<b>DAFTAR PUSTAKA .....</b>	<b>67</b>



## DAFTAR TABEL

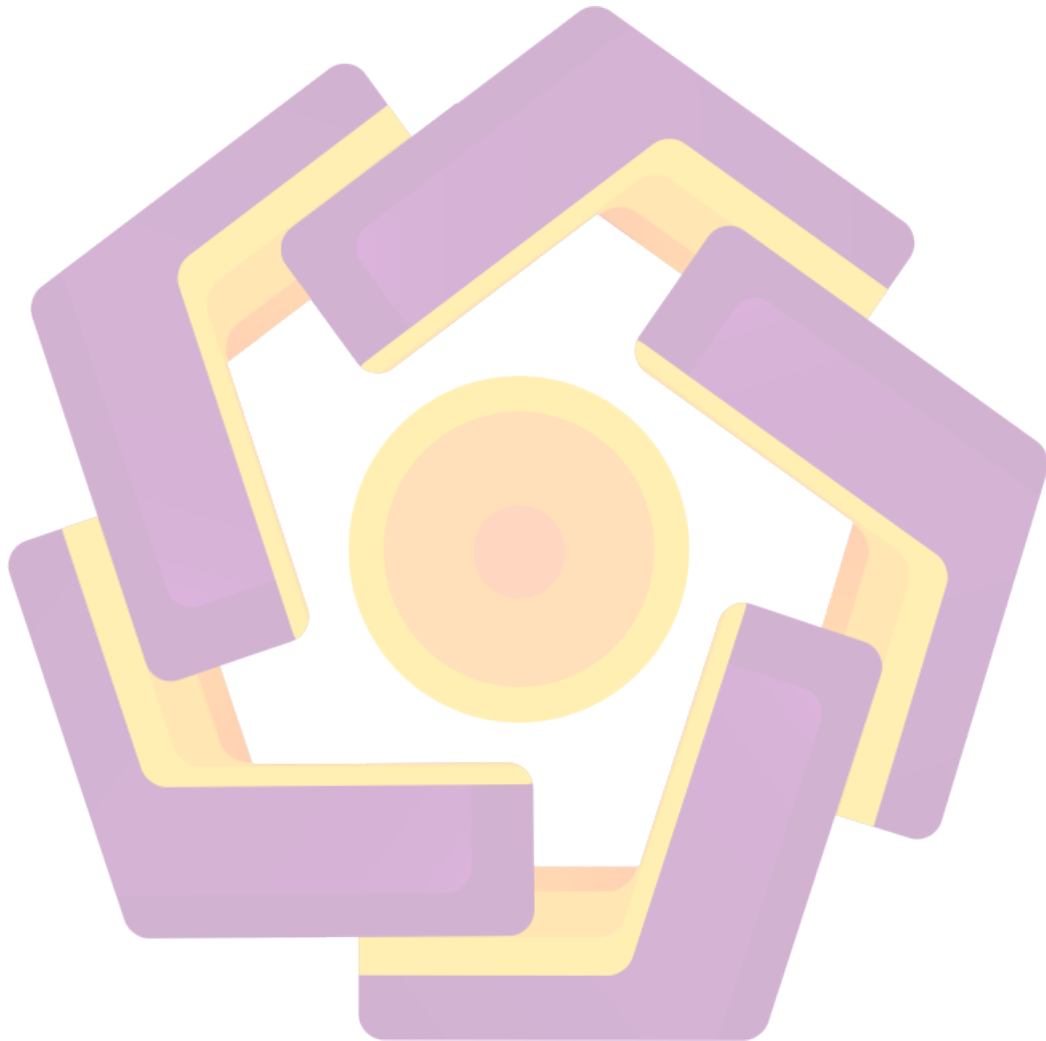
Tabel 2. 1 Perbandingan Penelitian Terkait Dengan Sistem Yang Dibangun .....	10
Tabel 3. 1 Data Tweet Sebelum Dilakukan Cleaning Data .....	23
Tabel 3. 2 Data Tweet Setelah Dilakukan Cleaning Data.....	25
Tabel 3. 3 Data Tweet Setelah <i>Case Folding</i> .....	25
Tabel 3. 4 Hasil <i>Tokenization</i> Dari <i>Case Foldin</i> .....	26
Tabel 3. 5 Data Setelah Dilakukan <i>Stopword &amp; Stemming</i> .....	27
Tabel 3. 6 Feature List .....	29
Tabel 3. 7 TF-IDF Positif.....	30
Tabel 3. 8 TF-IDF Negatif .....	31
Tabel 3. 9 Hasil <i>Cosine Similarity</i> Positif.....	35
Tabel 3. 10 Hasil <i>Cosine Similarity</i> Negatif .....	36
Tabel 3. 11 Hasil <i>Cosine Similarity</i> .....	37
Tabel 3. 12 Persamaan $w_1$ , $w_2$ , dan $b$ Pada Setiap Label .....	39
Tabel 3. 13 Perhitungan Persamaan Positif Terhadap Training Data Postif.....	40
Tabel 3. 14 Perhitungan Persamaan Negatif Terhadap Training Data Positif .....	40
Tabel 3. 15 Predik Pada Data Training .....	40
Tabel 3. 16 <i>Confusion Matrix</i> Data Sampel.....	41
Tabel 4. 1 <i>Confusion Matrix</i> SVM.....	60
Tabel 4. 2 Hasil 10 <i>Fold Cross Validation</i> SVM.....	61

## DAFTAR GAMBAR

Gambar 2. 1 Hyperplane pada <i>Support Vector Machine</i> .....	15
Gambar 2. 2 Pemisahan Dua Kelas Dengan Margin Maksimum .....	16
Gambar 2. 3 Contoh Confusion Matrix.....	18
Gambar 3. 1 Diagram Alir Preprocessing .....	43
Gambar 3. 2 Proses Klasifikasi Support Vector Machine.....	44
Gambar 3. 3 Konfigurasi Autentikasi File Crawling .....	45
Gambar 3. 4 Isi Pada File Training .....	46
Gambar 3. 5 Langkah-Langkah Klasifikasi .....	46
Gambar 3. 6 Langkah -Langkah Training Data .....	47
Gambar 3. 7 Langkah-Langkah Testing Data .....	49
Gambar 4. 1 <i>Script Crawling</i> Data Twitter .....	51
Gambar 4. 2 Hasil <i>Crawling</i> Data.....	52
Gambar 4. 3 <i>Dataset Training</i> yang akan digunakan .....	53
Gambar 4. 4 Script Akses File Data Training .....	54
Gambar 4. 5 <i>Script Cleaning</i> dan <i>Case Folding</i> .....	55
Gambar 4. 6 <i>Script Tokenization</i> .....	56
Gambar 4. 7 <i>Script Stopword</i> .....	56
Gambar 4. 8 Script Pembobotan dan Klasifikasi .....	57
Gambar 4. 9 Script Prediksi Sentimen .....	57
Gambar 4. 10 <i>Confusion Matrix SVM</i> .....	58

Gambar 4. 11 Hasil Prediksi dalam bentuk *Pie Chart* .....59

Gambar 4. 12 Grafik Performa Model Hasil *10 Fold Cross Validation* .....65





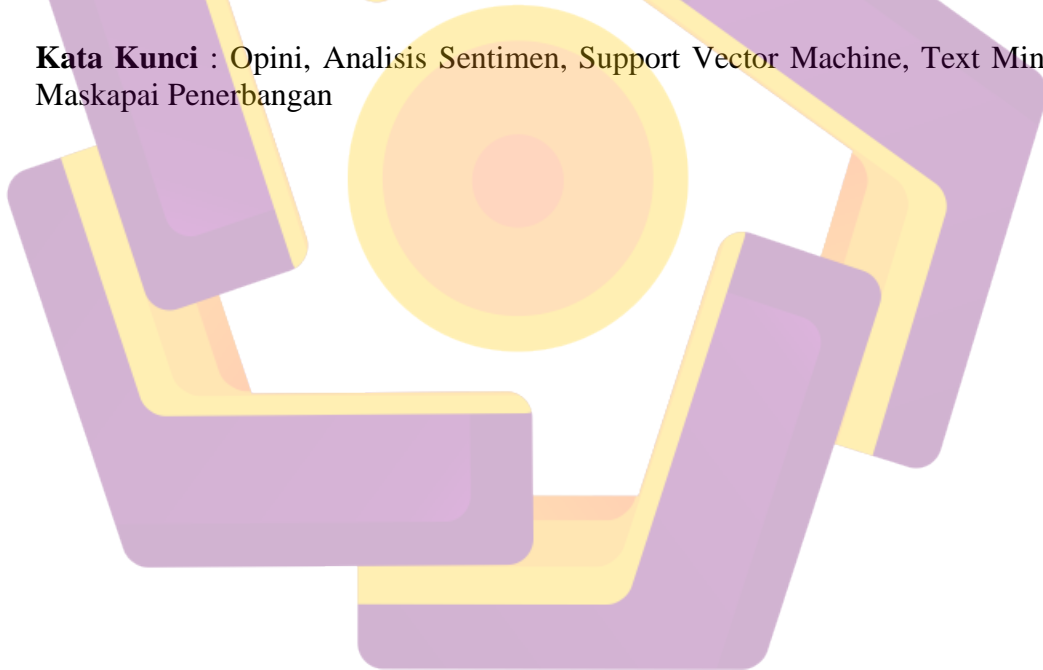
## INTISARI

Twitter merupakan salah satu media sosial yang digunakan untuk menyalurkan opini dan membahas berbagai topik yang ada di sekitar. Salah satu topik yang sering dibahas yaitu tentang maskapai penerbangan.

Pada saat ini transportasi udara menjadi sarana prasarana yang memudahkan masyarakat dalam berpergian dan maskapai penerbangan memegang peranan penting dalam hal transportasi udara. Oleh karena itu diperlukan suatu metode yang dapat digunakan untuk mengetahui opini masyarakat tentang maskapai penerbangan pada sosial media Twitter. Untuk menyelesaikan masalah tersebut maka diperlukan suatu metode yang dapat mengkategorikan opini tersebut.

*Support Vector Machine* merupakan salah satu metode dalam *Text Mining* yang dapat mengkategorikan opini-opini tersebut. Data yang diperoleh dari Twitter akan dilabeli dan dianalisis menggunakan metode SVM untuk mengklasifikasikan opini tersebut. Hasil dari klasifikasi menggunakan metode SVM diperoleh tingkat akurasi tertinggi sebesar 87%.

**Kata Kunci** : Opini, Analisis Sentimen, Support Vector Machine, Text Mining, Maskapai Penerbangan



## **ABSTRACT**

*Twitter is one of the social media that is used to express an opinions and discuss various topics around. One of the topics that is often discussed is about airlines.*

*At this time air transportation has become an infrastructure that makes it easier for people to travel and airlines play an important role in air transportation. Therefore we need a method that can be used to find out public opinion about airlines on social media Twitter. To solve this problem, we need a method that can categorize these opinions.*

*Support Vector Machine is one of the methods in Text Mining that can categorize these opinions. The data obtained from Twitter will be labeled and analyzed using the SVM method to classify these opinions. The results of the classification using the SVM method obtained the highest accuracy rate of 87%.*

**Keyword** : *Opinion, Sentiment Analysis, Support Vector Machine, Text Mining, Airlines*

