

**PERBANDINGAN KINERJA ALGORITMA  
*NAIVE BAYES* DAN *SUPPORT VECTOR MACHINE*  
PADA ANALISIS SENTIMEN  
MASKAPAI PENERBANGAN INDONESIA**

**(Skripsi)**

**Oleh**

**Muhammad Dzaki Arrahman**

**2017051073**



**JURUSAN ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
2024**

## ABSTRAK

### PERBANDINGAN KINERJA ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN MASKAPAI PENERBANGAN INDONESIA

Oleh

**Muhammad Dzaki Arrahman**

Meningkatnya integrasi teknologi ke dalam kehidupan sehari-hari telah menumbuhkan gaya hidup digital di mana platform media sosial memainkan peran penting dalam menyebarkan informasi. media sosial *X* (sebelumnya bernama *Twitter*), banyak digunakan untuk berbagi opini, berita, dan tren. Banyak topik populer yang dibahas di media sosial *X* adalah topik-topik yang berkaitan dengan industri penerbangan, khususnya maskapai penerbangan Indonesia. Penelitian ini melakukan analisis sentimen terhadap tweet yang berkaitan dengan empat maskapai penerbangan besar di Indonesia: Garuda, Lion Air, Batik Air, dan Sriwijaya Air dengan menggunakan algoritma *Naive Bayes* dan *Support Vector Machine*. Penelitian ini dilakukan pengumpulan data yang berfokus pada tweet yang berhubungan dengan maskapai penerbangan Indonesia dan memberikan label pada data ke dalam kategori positif, negatif, dan netral. Preprocessing diterapkan, yang melibatkan tugas-tugas seperti menghilangkan noise, tokenisasi, dan normalisasi. Untuk menyeimbangkan dataset, teknik augmentasi data diimplementasikan. Semua model pendekatan pembelajaran mesin dilatih, dan kinerjanya dievaluasi berdasarkan akurasi, presisi, recall, dan metrik F1-Score. Hasilnya menunjukkan bahwa SVM lebih unggul dengan akurasi sebesar 92% dibandingkan dengan *Naive Bayes* sebesar 89%. Hal ini menunjukkan bahwa SVM adalah model yang paling efektif untuk klasifikasi sentimen dengan benar dibandingkan dengan algoritma *Naive Bayes*.

Kata Kunci: Analisis Sentimen, *Support Vector Machine*, Maskapai Penerbangan Indonesia, *Naive Bayes*, *X*.

## **ABSTRACT**

### **A COMPARATIVE PERFORMANCE OF NAIVE BAYES AND SUPPORT VECTOR MACHINE IN SENTIMENT ANALYSIS OF INDONESIAN AIRLINES**

**By**

**Muhammad Dzaki Arrahman**

The growing integration of technology into everyday life has created a digital lifestyle where social media plays a key role in spreading information. X (formerly Twitter) is a popular platform for sharing opinions, news, and trends. A frequently discussed topic on X is the aviation industry, particularly Indonesian airlines. This study analyzes the sentiment of tweets related to four major Indonesian airlines: Garuda, Lion Air, Batik Air, and Sriwijaya Air, using two algorithms: Naive Bayes and Support Vector Machine (SVM). The research involved collecting tweets about these airlines and categorizing them as positive, negative, or neutral. After gathering the data, preprocessing was performed, including noise removal, tokenization, and normalization. To balance the dataset, data augmentation techniques were applied. Both machine learning models were trained, and their performance was measured using accuracy, precision, recall, and F1-Score. The results showed that SVM performed better, achieving 92% accuracy compared to Naive Bayes' 89%. This suggests that SVM is the more effective model for sentiment classification in this context.

**Keywords:** Indonesian Airlines, Sentiment Analysis, Support Vector Machine (SVM), Naive Bayes, X.

**PERBANDINGAN KINERJA ALGORITMA NAIVE BAYES DAN  
SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN  
MASKAPAI PENERBANGAN INDONESIA**

**Oleh**

**Muhammad Dzaki Arrahman**

**2017051073**

**Skripsi**

**Sebagai Salah Satu Syarat untuk Mencapai Gelar**

**SARJANA KOMPUTER**

**Pada**

**Jurusan Ilmu Komputer**

**Fakultas Matematika dan Ilmu Pengetahuan Alam**

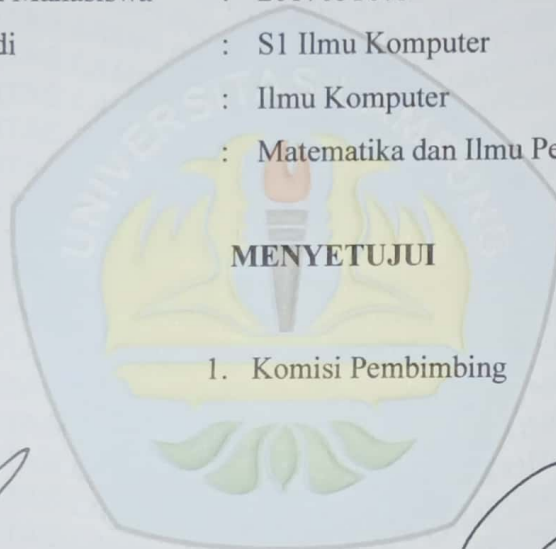
**Universitas Lampung**



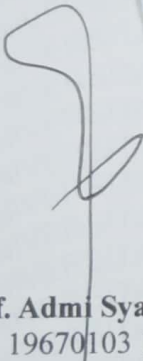
**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2024**

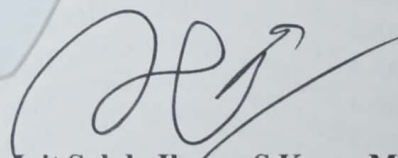
Judul Skripsi : **PERBANDINGAN KINERJA ALGORITMA NAIVE BAYES DAN SUPPORT VECTOR MACHINE PADA ANALISIS SENTIMEN MASKAPAI PENERBANGAN INDONESIA**

Nama Mahasiswa : **Muhammad Dzaki Arrahman**  
Nomor Induk Mahasiswa : 2017051073  
Program Studi : S1 Ilmu Komputer  
Jurusan : Ilmu Komputer  
Fakultas : Matematika dan Ilmu Pengetahuan Alam

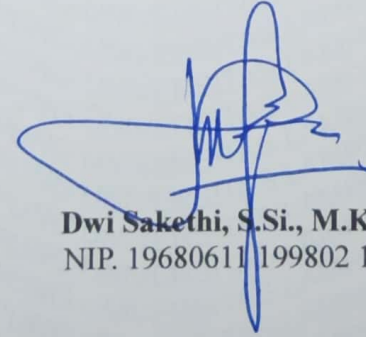


1. Komisi Pembimbing

  
**Prof. Admi Syarif, Ph.D**  
NIP. 19670103 199203 1 003

  
**Igit Sabda Ilman, S.Kom., M.Cs**  
NIP. 232111960101101

2. Ketua Jurusan Ilmu Komputer

  
**Dwi Sakethi, S.Si., M.Kom**  
NIP. 19680611 199802 1 001

**MENGESAHKAN**

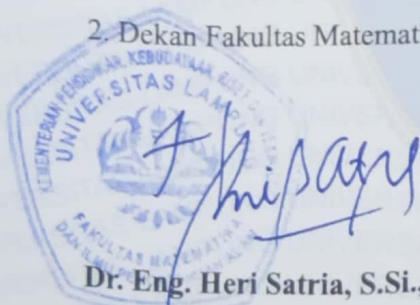
1. Tim Penguji

Ketua : **Prof. Admi Syarif, Ph.D**

Penguji I  
Sekretaris : **Igit Sabda Ilman, S.Kom., M.Cs**

Penguji II  
Pembahas : **Bambang Hermanto, S.Kom., M.Cs**

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



**Dr. Eng. Heri Satria, S.Si., M.Si**

NIP. 19711001 200501 1 002

Tanggal Lulus Ujian Skripsi : **08 Oktober 2024**

## PERNYATAAN

Saya yang bertanda tangan dibawah ini:

Nama : Muhammad Dzaki Arrahman

NPM : 2017051073

Menyatakan bahwa skripsi saya yang berjudul **“Perbandingan Kinerja Algoritma Naive Bayes Dan Support Vector Machine Pada Analisis Sentimen Maskapai Penerbangan Indonesia”** merupakan karya saya sendiri dan bukan karya orang lain. Semua tulisan yang tertuang di skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila di kemudian hari terbukti skripsi saya merupakan hasil penjiplakan atau dibuat orang lain, maka saya bersedia menerima sanksi berupa pencabutan gelar yang telah saya terima.

Bandar Lampung, 06 Desember 2024



Muhammad Dzaki Arrahman  
NPM. 2017051073

## RIWAYAT HIDUP



Lahir di Kota DKI Jakarta pada tanggal 05 Oktober 2002. Anak pertama dari dua bersaudara, dari Bapak Raeno dan Ibu Qomariah. Menyelesaikan Pendidikan di SDIT Al-Azkar pada Tahun 2014, kemudian menyelesaikan pendidikan menengah pertama di MTS Negeri 1 Tangerang Selatan pada Tahun 2017, dan lulus dari pendidikan menengah atas di SMA Negeri 3 Tangerang Selatan pada Tahun 2020.

Pada tahun 2020, terdaftar sebagai mahasiswa Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SBMPTN. Kegiatan-kegiatan yang dilakukan selama menjadi mahasiswa adalah sebagai berikut.

1. Menjadi anggota Bidang Informasi dan Komunikasi ROIS FMIPA Universitas Lampung pada Tahun 2021
2. Menjadi Koordinator pada kegiatan Festival Islam FMIPA Bidang Publikasi Dekorasi dan Dokumentasi ROIS FMIPA pada Tahun 2021
3. Mengikuti program beasiswa pelatihan pemrogram web yang dilaksanakan oleh PT. Udacoding Indonesia pada Tahun 2021



4. Menjadi anggota Bidang Informasi dan Komunikasi ROIS FMIPA Universitas Lampung Pada Tahun 2022.
5. Menjadi anggota pada kegiatan Festival Islam FMIPA Bidang Publikasi Dekorasi dan Dokumentasi ROIS FMIPA pada Tahun 2022.
6. Melaksanakan Kerja Praktik pada Tahun 2022 di PT Telkom Witel Lampung Indonesia.
7. Melaksanakan kegiatan pelatihan keterampilan bidang data analisis yang dilaksanakan oleh REVOU pada Tahun 2023.
8. Melaksanakan kegiatan Magang Merdeka Belajar – Kampus Merdeka (MBKM) pada Tahun 2023 di Sekolah Kak Seto.
9. Melaksanakan Kuliah Kerja Nyata di Sekolah Kak Seto pada Tahun 2023 di Kecamatan Pondok Aren, Kota Tangerang.
10. Melaksanakan kegiatan pelatihan data science, IT Fundamentals, dan Cyber Security yang dilaksanakan oleh KOMINFO pada Tahun 2024.

## **MOTTO**

“Ketika kamu mengagumi langit, kamu juga harus menyukai mendungnya,  
hujannya, petirnya, tidak hanya saat ia menampakkan Pelangi dan senjanya.”

**(Khalil Gibran)**

“Permulaan dari semua pencapaian adalah keinginan.”

**(Napoleon Hill)**

“Pendidikan adalah senjata paling ampuh yang bisa digunakan untuk mengubah  
dunia.”

**(Nelson Mandela)**

“Sukses datang kepada mereka yang mampu membangun jaringan yang kuat dan  
memeliharanya dengan baik.”

**(Keith Ferrazzi)**

## **PERSEMBAHAN**

Puji dan syukur tercurahkan kepada Allah *Subhanahu Wa Ta'ala* atas segala Rahmat dan Karunia-Nya sehingga saya dapat menyelesaikan skripsi ini.

Sholawat serta salam selalu tercurahkan kepada Nabi Muhammad *Sholallahu*

*Alaihi Wasallam.*

Kupersembahkan karya ini kepada :

### **Kedua Orangtua, Adik, dan Keluarga Besar**

Yang senantiasa memberikan yang terbaik, dan melantunkan do'a yang selalu menyertaiku. Kuucapkan juga terimakasih sebesar-besarnya karena telah mendidik dan membesarkanku dengan cara yang dipenuhi kasih sayang, dukungan, dan pengorbanan yang luar biasa.

### **Seluruh Keluarga Besar Ilmu Komputer 2020**

Yang selalu memberikan semangat dan dukungan.

### **Almamater tercinta, Universitas Lampung dan Jurusan Ilmu Komputer**

Tempat menimba ilmu, untuk menjadi bekal hidup dunia dan akhirat.

## SANWACANA

Puji Syukur kehadiran Allah *Subhanahu Wa Ta'ala* atas berkah, rahmat dan hidayatNya, serta petunjuk dan pedoman dari Rasulullah Nabi Muhammad *Sholallahu Alaihi Wasallam* penulis dapat menyelesaikan skripsi yang berjudul “Perbandingan Kinerja Algoritma *Naive Bayes* Dan *Support Vector Machine* Pada Analisis Sentimen Maskapai Penerbangan Indonesia” dengan baik.

Terima kasih penulis ucapkan kepada semua pihak yang telah membantu dan berperan besar dalam menyusun skripsi ini, antara lain:

1. Kedua orangtua dan adik tercinta yang memberikan dukungan, do'a, semangat, motivasi, dan kasih sayang yang luar biasa tak terhingga. Semua yang telah kalian berikan tidak akan pernah mampu untuk bisa dibalas. Semoga Allah Subhanahu Wa Ta'ala selalu memberikan kebahagiaan dan keberkahan dalam kehidupan kalian di dunia dan akhirat.
2. Bapak Dr. Eng. Heri Satria, S.Si., M.Si. selaku Dekan FMIPA Universitas Lampung.
3. Bapak Dwi Sakethi, S.Kom., M.Kom selaku Ketua Jurusan Ilmu Komputer Universitas Lampung.
4. Ibu Anie Rose Irawati S.T., M.Cs. selaku Sekretaris Jurusan Ilmu Komputer FMIPA Universitas Lampung.
5. Prof. Admi Syarif, Ph.D sebagai dosen Pembimbing Utama yang telah memberikan arahan, ide, motivasi, dan dukungan selama proses skripsi ini.
6. Bapak Igit Sabda Iلمان, S.Kom., M.Cs sebagai dosen Pembimbing Pembantu yang telah memberikan arahan, ide, motivasi, saran selama proses skripsi ini.

7. Bapak Bambang Hermanto, S.Kom., M.Cs sebagai dosen Pembahas yang telah memberikan masukan untuk skripsi yang dibuat.
8. Ibu Ossy Dwiendah Wulansari, S.Si., M.T sebagai dosen Pembimbing Akademik yang telah mendukung kegiatan akademik penulis selama perkuliahan.
9. Almarhum Bapak Irwan Adi Pribadi sebagai dosen pembimbing akademik yang telah mendukung kegiatan akademik penulis selama perkuliahan.
10. Bapak dan Ibu Dosen Jurusan Ilmu Komputer FMIPA Universitas Lampung yang telah memberikan ilmu dan pengalaman selama perkuliahan untuk menjadi lebih baik.
11. Ibu Ade Nora Maela, Bang Zainuddin, Mas Nofal, dan Mas Sam yang telah membantu segala urusan administrasi penulis dan kebutuhan lainnya di Jurusan Ilmu Komputer.
12. Teman-teman satu bimbingan Prof Admi (Pak Yodhi, Pak Ari, Pak Nurjoko, Pak Riko, Pak Nizar, Ibu Apri, Joy, Safiira, Yoan, dan Nabila) yang menemani penulis dan membantu penulis selama proses skripsi.
13. Keluarga besar Ilmu Komputer 2020 yang tidak bisa penulis sebutkan satu persatu.

Bandar Lampung, 06 Desember 2024

Muhammad Dzaki Arrahman  
NPM. 2017051073

## DAFTAR ISI

DAFTAR ISI .....	xiii
DAFTAR GAMBAR .....	xv
DAFTAR TABEL .....	xix
<b>I. PENDAHULUAN</b> .....	1
<b>1.1 Latar Belakang</b> .....	1
<b>1.2 Rumusan Masalah</b> .....	6
<b>1.3 Batasan Masalah</b> .....	6
<b>1.4 Tujuan Penelitian</b> .....	6
<b>1.5 Manfaat Penelitian</b> .....	7
<b>II. TINJAUAN PUSTAKA</b> .....	8
<b>2.1 Maskapai Penerbangan</b> .....	8
<b>2.2 <i>Text Mining</i></b> .....	9
<b>2.3 Analisis Sentimen</b> .....	10
<b>2.4 Media Sosial X</b> .....	10
<b>2.5 <i>K-Fold cross validation</i></b> .....	11
<b>2.6 <i>Naive Bayes</i></b> .....	12
<b>2.7 <i>Support Vector Machine</i></b> .....	14
<b>2.8 <i>Confusion Matrix</i></b> .....	16
<b>2.9 <i>Python</i></b> .....	16
<b>2.8. Penelitian Terdahulu</b> .....	18
<b>III. METODE PENELITIAN</b> .....	20
<b>3.1 Waktu dan Tempat Penelitian</b> .....	20
<b>3.2 Perangkat Penelitian</b> .....	20
<b>3.3 Tahapan Penelitian</b> .....	21
<b>I. Pengumpulan Data</b> .....	22
<b>II. Pelabelan Data</b> .....	22

III.	<i>Preprocessing Data</i> .....	22
IV.	Augmentasi Data .....	24
V.	Pembagian Data .....	25
VI.	Evaluasi .....	25
VII.	Penarikan Kesimpulan .....	27
IV.	HASIL DAN PEMBAHASAN .....	28
4.1	Dataset .....	28
I.	Maskapai Indonesia .....	28
II.	Pelabelan Data .....	34
III.	<i>Preprocessing Data</i> .....	48
IV.	Augmentasi Data .....	53
V.	Pembagian Data .....	55
4.2	Klasifikasi Model <i>Naive Bayes</i> dan <i>Support Vector Machine</i> .....	56
I.	Klasifikasi metode <i>Naive Bayes</i> .....	56
II.	Klasifikasi Metode <i>Support Vector Machine</i> .....	95
III.	Perbandingan Model <i>Naive Bayes</i> dan <i>Support Vector Machine</i> .....	134
V.	KESIMPULAN DAN SARAN .....	166
5.1	Kesimpulan .....	166
5.2	Saran .....	169

## DAFTAR GAMBAR

### Halaman

Gambar 1. Ilustrasi Support Vector Machine. ....	15
Gambar 2. Tahapan Penelitian. ....	21
Gambar 3. Alur preprocessing data. ....	23
Gambar 4. Proses pengambilan data pada media sosial X. ....	28
Gambar 5. Perbandingan data sentimen positif, negatif, dan netral. ....	36
Gambar 6. Sebaran data sentimen positif, negatif, dan netral di media sosial X. ....	37
Gambar 7. Sentimen masyarakat pengguna media sosial X terhadap maskapai Garuda tahun 2023. ....	39
Gambar 8. Persentase sentimen masyarakat pengguna media sosial X terhadap maskapai Garuda tahun 2023. ....	40
Gambar 9. Sentimen masyarakat pengguna media sosial X terhadap maskapai Lion Air. ....	42
Gambar 10. Persentase sentimen masyarakat pengguna media sosial X terhadap maskapai Lion Air. ....	42
Gambar 11. Sentimen masyarakat pengguna media sosial X terhadap maskapai Batik Air. ....	44
Gambar 12. Persentase sentimen masyarakat pengguna media sosial X terhadap maskapai Batik Air. ....	45
Gambar 13. Sentimen masyarakat pengguna media sosial X terhadap maskapai Sriwijaya Air. ....	47
Gambar 14. Persentase sentimen masyarakat pengguna media sosial X terhadap maskapai Sriwijaya Air. ....	47
Gambar 15. Proses <i>cleaning</i> pada <i>python</i> - tahap 1. ....	49
Gambar 16. Proses <i>cleaning</i> pada <i>python</i> - tahap 2. ....	49



Gambar 17. Proses <i>cleaning</i> pada <i>python</i> - tahap 3 .....	50
Gambar 18. Proses <i>Case Folding</i> dan <i>Tokenizing</i> pada <i>python</i> . .....	51
Gambar 19. Proses <i>removing stopwords</i> dalam <i>python</i> . .....	52
Gambar 20. Proses <i>stemming</i> data dalam <i>python</i> . .....	53
Gambar 21. Proses pembagian data dalam <i>python</i> . .....	56
Gambar 22. Penerapan 10-fold crossvalidation pada dataset maskapai Garuda metode <i>Naive Bayes</i> .....	56
Gambar 23. Penerapan 10-fold cross validation pada dataset maskapai Lion Air metode <i>Naive Bayes</i> .....	64
Gambar 24. Penerapan 10-fold cross validation pada dataset maskapai Batik Air metode <i>Naive Bayes</i> .....	71
Gambar 25. Penerapan 10-Fold cross validation pada Dataset Maskapai Sriwijaya Air metode <i>Naive Bayes</i> . .....	79
Gambar 26. Penerapan 10-Fold cross validation pada Dataset Maskapai Indonesia metode <i>Naive Bayes</i> . .....	87
Gambar 27. Penerapan 10-fold cross validation pada dataset maskapai Garuda metode <i>Support Vector Machine</i> . .....	95
Gambar 28. Penerapan 10-fold cross validation pada dataset maskapai Lion Air metode <i>Support Vector Machine</i> . .....	103
Gambar 29. Penerapan 10-fold cross validation pada dataset maskapai Batik Air metode <i>Support Vector Machine</i> . .....	111
Gambar 30. Penerapan 10-Fold cross validation pada Dataset Maskapai Sriwijaya Air metode <i>Support Vector Machine</i> . .....	119
Gambar 31. Penerapan 10-Fold cross validation pada Dataset Maskapai Indonesia metode <i>Support Vector Machine</i> . .....	127
Gambar 32. Perbandingan Rata-rata Akurasi Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Garuda. ....	136
Gambar 33. Perbandingan Rata-rata <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Garuda. ....	137
Gambar 34. Perbandingan Rata-rata <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> . .....	139

Gambar 35. Perbandingan Rata-rata F1-Score Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Garuda.....	140
Gambar 36. Perbandingan Rata-rata Akurasi Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Lion Air. ....	142
Gambar 37. Perbandingan Rata-rata <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Lion Air. ....	143
Gambar 38. Perbandingan Rata-rata <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Lion Air. ....	145
Gambar 39. Perbandingan Rata-rata F1-Score Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Lion Air. ....	147
Gambar 40. Perbandingan Rata-rata Akurasi Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Batik Air. ....	148
Gambar 41. Perbandingan Rata-rata <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Batik Air. ....	149
Gambar 42. Perbandingan Rata-rata <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Batik Air. ....	151
Gambar 43. Perbandingan Rata-rata F1-Score Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Batik Air. ....	152
Gambar 44. Perbandingan Rata-rata Akurasi Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Sriwijaya Air.....	154
Gambar 45. Perbandingan Rata-rata <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Sriwijaya Air.....	155
Gambar 46. Perbandingan Rata-rata <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Sriwijaya Air.....	157
Gambar 47. Perbandingan Rata-rata F1-Score Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Sriwijaya Air.....	159
Gambar 48. Perbandingan Rata-rata Akurasi Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> Maskapai Penerbangan Indonesia.....	160
Gambar 49. Perbandingan Rata-rata <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Penerbangan Indonesia. ....	162
Gambar 50. Perbandingan Rata-rata <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Penerbangan Indonesia. ....	164

Gambar 51. Perbandingan Rata-rata F1-Score Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Penerbangan Indonesia. ....	165
--	-----

## DAFTAR TABEL

	Halaman
Tabel 1. Jumlah Pesawat Milik Maskapai Penerbangan Indonesia.....	8
Tabel 2. 10 Aplikasi Dengan Pengguna Terbanyak di Indonesia.....	11
Tabel 3. Ilustasi pembagian data menggunakan <i>k-fold cross validation</i> .....	12
Tabel 4. Penerapan Confusion Matrix untuk evaluasi kinerja model .....	16
Tabel 5. Penelitian Terdahulu.....	18
Tabel 6. Pembagian data latih dan data uji.....	25
Tabel 7. Confusion Matrix .....	26
Tabel 8. Maskapai penerbangan Indonesia .....	29
Tabel 9. Data maskapai Garuda tahun 2023.....	30
Tabel 10. Data maskapai Lion <i>Air</i> tahun 2023.....	31
Tabel 11. Data maskapai Batik <i>Air</i> tahun 2023.....	32
Tabel 12. Data maskapai Sriwijaya <i>Air</i> tahun 2023 .....	33
Tabel 13. Pelabelan Data sentimen positif oleh tiga anator .....	34
Tabel 14. Pelabelan Data sentimen negatif oleh tiga anator .....	35
Tabel 15. Pelabelan Data sentimen netral oleh tiga anator .....	35
Tabel 16. Data media sosial X yang sudah diberi label .....	36
Tabel 17. Tabel sentimen positif, negatif, dan netral.....	37
Tabel 18. Data hasil pelabelan data pada maskapai Garuda .....	38
Tabel 19. Data hasil pelabelan data pada maskapai Lion <i>Air</i> .....	41
Tabel 20. Data hasil pelabelan data pada maskapai Batik <i>Air</i> .....	43
Tabel 21. Data hasil pelabelan data pada maskapai Sriwijaya <i>Air</i> .....	46
Tabel 22. Penerapan <i>cleaning</i> pada dataset.....	48
Tabel 23. Penerapan <i>Case Folding</i> dan tokenizing pada dataset .....	50
Tabel 24. Penerapan stopword removal pada dataset.....	51

Tabel 25. Penerapan <i>stemming</i> pada dataset .....	52
Tabel 26. Penerapan Augmentasi Data pada Maskapai Garuda.....	53
Tabel 27. Penerapan Augmentasi Data pada Maskapai Lion <i>Air</i> .....	54
Tabel 28. Penerapan Augmentasi Data pada Maskapai Batik <i>Air</i> .....	54
Tabel 29. Penerapan Augmentasi Data pada Maskapai Sriwijaya <i>Air</i> .....	55
Tabel 30. Dataset yang sudah dibagi menjadi data latih dan data uji .....	55
Tabel 31. Hasil Akurasi pada Masing-Masing Fold Metode <i>Naive Bayes</i> pada Garuda.....	57
Tabel 32. Confusion Matrix Fold-1 Maskapai Garuda Metode <i>Naive Bayes</i> .....	57
Tabel 33. Confusion Matrix Fold-2 Maskapai Garuda Metode <i>Naive Bayes</i> .....	58
Tabel 34. Confusion Matrix Fold-3 Maskapai Garuda Metode <i>Naive Bayes</i> .....	58
Tabel 35. Confusion Matrix Fold-4 Maskapai Garuda Metode <i>Naive Bayes</i> .....	59
Tabel 36. Confusion Matrix Fold-5 Maskapai Garuda Metode <i>Naive Bayes</i> .....	59
Tabel 37. Confusion Matrix Fold-6 Maskapai Garuda Metode <i>Naive Bayes</i> .....	60
Tabel 38. Confusion Matrix Fold-7 Maskapai Garuda Metode <i>Naive Bayes</i> .....	60
Tabel 39. Confusion Matrix Fold-8 Maskapai Garuda Metode <i>Naive Bayes</i> .....	61
Tabel 40. Confusion Matrix Fold-9 Maskapai Garuda Metode <i>Naive Bayes</i> .....	61
Tabel 41. Confusion Matrix Fold-10 Maskapai Garuda Metode <i>Naive Bayes</i> .....	62
Tabel 42. Nilai TP,FP,FN,TN masing-masing fold metode <i>Naive Bayes</i> pada maskapai Garuda.....	62
Tabel 43. Pengukuran kinerja <i>Naive Bayes</i> pada dataset maskapai Garuda .....	63
Tabel 45. Hasil akurasi masing-masing fold metode <i>Naive Bayes</i> pada dataset maskapai Lion <i>Air</i> .....	64
Tabel 46. Confusion Matrix Fold-1 Maskapai Lion <i>Air</i> Metode <i>Naive Bayes</i> .....	65
Tabel 47. Confusion Matrix Fold-2 Maskapai Lion <i>Air</i> Metode <i>Naive Bayes</i> .....	65
Tabel 48. Confusion Matrix Fold-3 Maskapai Lion <i>Air</i> Metode <i>Naive Bayes</i> .....	66
Tabel 49. Confusion Matrix Fold-4 Maskapai Lion <i>Air</i> Metode <i>Naive Bayes</i> .....	66
Tabel 50. Confusion Matrix Fold-5 Maskapai Lion <i>Air</i> Metode <i>Naive Bayes</i> .....	67
Tabel 51. Confusion Matrix Fold-6 Maskapai Lion <i>Air</i> Metode <i>Naive Bayes</i> .....	67
Tabel 52. Confusion Matrix Fold-7 Maskapai Lion <i>Air</i> Metode <i>Naive Bayes</i> .....	68
Tabel 53. Confusion Matrix Fold-8 Maskapai Lion <i>Air</i> Metode <i>Naive Bayes</i> .....	68
Tabel 54. Confusion Matrix Fold-9 Maskapai Lion <i>Air</i> Metode <i>Naive Bayes</i> .....	69

Tabel 55. Confusion Matrix Fold-10 Maskapai Lion <i>Air</i> Metode <i>Naive Bayes</i> ...	69
Tabel 56. Nilai TP,FP,FN,TN masing-masing fold metode <i>Naive Bayes</i> pada maskapai Lion <i>Air</i> .....	70
Tabel 57. Pengukuran kinerja <i>Naive Bayes</i> pada dataset maskapai Lion <i>Air</i> .....	70
Tabel 58. Lanjutan Tabel 57 .....	71
Tabel 59. Hasil akurasi masing-masing fold metode <i>Naive Bayes</i> pada dataset maskapai Lion <i>Air</i> .....	72
Tabel 60. Confusion Matrix Fold-1 Maskapai Batik <i>Air</i> Metode <i>Naive Bayes</i> ....	72
Tabel 61. Confusion Matrix Fold-2 Maskapai Batik <i>Air</i> Metode <i>Naive Bayes</i> ....	73
Tabel 62. Confusion Matrix Fold-3 Maskapai Batik <i>Air</i> Metode <i>Naive Bayes</i> ....	73
Tabel 63. Confusion Matrix Fold-4 Maskapai Batik <i>Air</i> Metode <i>Naive Bayes</i> ....	74
Tabel 64. Confusion Matrix Fold-5 Maskapai Batik <i>Air</i> Metode <i>Naive Bayes</i> ....	74
Tabel 65. Confusion Matrix Fold-6 Maskapai Batik <i>Air</i> Metode <i>Naive Bayes</i> ....	75
Tabel 66. Confusion Matrix Fold-7 Maskapai Batik <i>Air</i> Metode <i>Naive Bayes</i> ....	75
Tabel 67. Confusion Matrix Fold-8 Maskapai Batik <i>Air</i> Metode <i>Naive Bayes</i> ....	76
Tabel 68. Confusion Matrix Fold-9 Maskapai Batik <i>Air</i> Metode <i>Naive Bayes</i> ....	76
Tabel 69. Confusion Matrix Fold-10 Maskapai Batik <i>Air</i> Metode <i>Naive Bayes</i> ..	77
Tabel 70. Nilai TP,FP,FN,TN masing-masing fold metode <i>Naive Bayes</i> pada maskapai Batik <i>Air</i> .....	77
Tabel 71. Pengukuran kinerja <i>Naive Bayes</i> pada dataset maskapai Batik <i>Air</i> .....	78
Tabel 72. Hasil akurasi masing-masing fold metode <i>Naive Bayes</i> pada dataset maskapai Sriwijaya <i>Air</i> .....	79
Tabel 73. Confusion Matrix Fold-1 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	80
Tabel 74. Confusion Matrix Fold-2 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	80
Tabel 75. Confusion Matrix Fold-3 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	81
Tabel 76. Confusion Matrix Fold-4 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	81
Tabel 77. Confusion Matrix Fold-5 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	82

Tabel 78. Confusion Matrix Fold-6 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	82
Tabel 79. Confusion Matrix Fold-7 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	83
Tabel 80. Confusion Matrix Fold-8 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	83
Tabel 81. Confusion Matrix Fold-9 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	84
Tabel 82. Confusion Matrix Fold-10 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	84
Tabel 83. Nilai TP,FP,FN,TN Masing-masing Fold Metode <i>Naive Bayes</i> pada Maskapai Sriwijaya <i>Air</i> .....	85
Tabel 84. Pengukuran kinerja <i>Naive Bayes</i> pada dataset maskapai Sriwijaya <i>Air</i>	86
Tabel 85. Hasil Akurasi Masing-Masing Fold Metode <i>Naive Bayes</i> Pada Dataset Maskapai Penerbangan Indonesia.....	87
Tabel 86. Confusion Matrix Fold-1 Maskapai Penerbangan Indonesia Metode <i>Naive Bayes</i> .....	88
Tabel 87. Confusion Matrix Fold-2 Maskapai Penerbangan Indonesia Metode <i>Naive Bayes</i> .....	88
Tabel 88. Confusion Matrix Fold-3 Maskapai Penerbangan Indonesia Metode <i>Naive Bayes</i> .....	89
Tabel 89. Confusion Matrix Fold-4 Maskapai Penerbangan Indonesia Metode <i>Naive Bayes</i> .....	89
Tabel 90. Confusion Matrix Fold-5 Maskapai Penerbangan Indonesia Metode <i>Naive Bayes</i> .....	90
Tabel 91. Confusion Matrix Fold-6 Maskapai Penerbangan Indonesia Metode <i>Naive Bayes</i> .....	90
Tabel 92. Confusion Matrix Fold-7 Maskapai Penerbangan Indonesia Metode <i>Naive Bayes</i> .....	91
Tabel 93. Confusion Matrix Fold-8 Maskapai Penerbangan Indonesia Metode <i>Naive Bayes</i> .....	91

Tabel 94. Confusion Matrix Fold-9 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	92
Tabel 95. Confusion Matrix Fold-10 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	92
Tabel 96. Nilai TP,FP,FN,TN Masing-masing Fold Metode <i>Naive Bayes</i> pada Maskapai Sriwijaya <i>Air</i> .....	93
Tabel 97. Pengukuran Kinerja <i>Naive Bayes</i> pada Dataset Maskapai Penerbangan Indonesia .....	94
Tabel 98. Hasil Akurasi pada Masing-Masing Fold Metode <i>Support Vector</i> <i>Machine</i> pada Garuda .....	96
Tabel 99. Confusion Matrix Fold-1 Maskapai Garuda Metode <i>Support Vector</i> <i>Machine</i> .....	96
Tabel 100. Confusion Matrix Fold-2 Maskapai Garuda Metode <i>Support Vector</i> <i>Machine</i> .....	97
Tabel 101. Confusion Matrix Fold-3 Maskapai Garuda Metode <i>Support Vector</i> <i>Machine</i> .....	97
Tabel 102. Confusion Matrix Fold-4 Maskapai Garuda Metode <i>Support Vector</i> <i>Machine</i> .....	98
Tabel 103. Confusion Matrix Fold-5 Maskapai Garuda Metode <i>Support Vector</i> <i>Machine</i> .....	98
Tabel 104. Confusion Matrix Fold-6 Maskapai Garuda Metode <i>Naive Bayes</i> .....	99
Tabel 105. Confusion Matrix Fold-7 Maskapai Garuda Metode <i>Naive Bayes</i> .....	99
Tabel 106. Confusion Matrix Fold-8 Maskapai Garuda Metode <i>Support Vector</i> <i>Machine</i> .....	100
Tabel 107. Confusion Matrix Fold-9 Maskapai Garuda Metode <i>Support Vector</i> <i>Machine</i> .....	100
Tabel 108. Confusion Matrix Fold-10 Maskapai Garuda Metode <i>Support Vector</i> <i>Machine</i> .....	101
Tabel 109. Nilai TP,FP,FN,TN masing-masing fold metode <i>Support Vector</i> <i>Machine</i> pada maskapai Garuda .....	101
Tabel 110. Pengukuran kinerja <i>Support Vector Machine</i> pada dataset maskapai Garuda.....	102



Tabel 111. Hasil akurasi masing-masing fold metode <i>Support Vector Machine</i> pada dataset maskapai <i>Lion Air</i> .....	103
Tabel 112. Confusion Matrix Fold-1 Maskapai <i>Lion Air</i> Metode <i>Support Vector Machine</i> .....	104
Tabel 113. Confusion Matrix Fold-2 Maskapai <i>Lion Air</i> Metode <i>Support Vector Machine</i> .....	104
Tabel 114. Confusion Matrix Fold-3 Maskapai <i>Lion Air</i> Metode <i>Support Vector Machine</i> .....	105
Tabel 115. Confusion Matrix Fold-4 Maskapai <i>Lion Air</i> Metode <i>Support Vector Machine</i> .....	105
Tabel 116. Confusion Matrix Fold-5 Maskapai <i>Lion Air</i> Metode <i>Support Vector Machine</i> .....	106
Tabel 117. Confusion Matrix Fold-6 Maskapai <i>Lion Air</i> Metode <i>Support Vector Machine</i> .....	106
Tabel 118. Confusion Matrix Fold-7 Maskapai <i>Lion Air</i> Metode <i>Support Vector Machine</i> .....	107
Tabel 119. Confusion Matrix Fold-8 Maskapai <i>Lion Air</i> Metode <i>Support Vector Machine</i> .....	107
Tabel 120. Confusion Matrix Fold-9 Maskapai <i>Lion Air</i> Metode <i>Support Vector Machine</i> .....	108
Tabel 121. Confusion Matrix Fold-10 Maskapai <i>Lion Air</i> Metode <i>Support Vector Machine</i> .....	108
Tabel 122. Nilai TP,FP,FN,TN masing-masing fold metode <i>Support Vector Machine</i> pada maskapai <i>Lion Air</i> .....	109
Tabel 123. Pengukuran kinerja <i>Support Vector Machine</i> pada dataset maskapai <i>Lion Air</i> .....	110
Tabel 124. Hasil akurasi masing-masing fold metode <i>Support Vector Machine</i> pada dataset maskapai <i>Batik Air</i> .....	111
Tabel 125. Confusion Matrix Fold-1 Maskapai <i>Batik Air</i> Metode <i>Support Vector Machine</i> .....	112
Tabel 126. Confusion Matrix Fold-2 Maskapai <i>Batik Air</i> Metode <i>Support Vector Machine</i> .....	112

Tabel 127. Confusion Matrix Fold-3 Maskapai Batik <i>Air</i> Metode <i>Support Vector Machine</i> .....	113
Tabel 128. Confusion Matrix Fold-4 Maskapai Batik <i>Air</i> Metode <i>Support Vector Machine</i> .....	113
Tabel 129. Confusion Matrix Fold-5 Maskapai Batik <i>Air</i> Metode <i>Support Vector Machine</i> .....	114
Tabel 130. Confusion Matrix Fold-6 Maskapai Batik <i>Air</i> Metode <i>Support Vector Machine</i> .....	114
Tabel 131. Confusion Matrix Fold-7 Maskapai Batik <i>Air</i> Metode <i>Support Vector Machine</i> .....	115
Tabel 132. Confusion Matrix Fold-8 Maskapai Batik <i>Air</i> Metode <i>Support Vector Machine</i> .....	115
Tabel 133. Confusion Matrix Fold-9 Maskapai Batik <i>Air</i> Metode <i>Support Vector Machine</i> .....	116
Tabel 134. Confusion Matrix Fold-10 Maskapai Batik <i>Air</i> Metode <i>Support Vector Machine</i> .....	116
Tabel 135. Nilai TP,FP,FN,TN masing-masing fold metode <i>Naive Bayes</i> pada maskapai Batik <i>Air</i> .....	117
Tabel 136. Pengukuran kinerja <i>Support Vector Machine</i> pada dataset maskapai Batik <i>Air</i> .....	118
Tabel 137. Hasil akurasi masing-masing fold metode <i>Support Vector Machine</i> pada dataset maskapai Sriwijaya <i>Air</i> .....	119
Tabel 138. Confusion Matrix Fold-1 Maskapai Sriwijaya <i>Air</i> Metode <i>Support Vector Machine</i> .....	120
Tabel 139. Confusion Matrix Fold-2 Maskapai Sriwijaya <i>Air</i> Metode <i>Support Vector Machine</i> .....	120
Tabel 140. Confusion Matrix Fold-3 Maskapai Garuda Metode <i>Support Vector Machine</i> .....	121
Tabel 141. Confusion Matrix Fold-4 Maskapai Sriwijaya <i>Air</i> Metode <i>Support Vector Machine</i> .....	121
Tabel 142. Confusion Matrix Fold-5 Maskapai Sriwijaya <i>Air</i> Metode <i>Support Vector Machine</i> .....	122

Tabel 143. Confusion Matrix Fold-6 Maskapai Sriwijaya <i>Air</i> Metode <i>Support Vector Machine</i> .....	122
Tabel 144. Confusion Matrix Fold-7 Maskapai Sriwijaya <i>Air</i> Metode <i>Support Vector Machine</i> .....	123
Tabel 145. Confusion Matrix Fold-8 Maskapai Sriwijaya <i>Air</i> Metode <i>Naive Bayes</i> .....	123
Tabel 146. Confusion Matrix Fold-9 Maskapai Sriwijaya <i>Air</i> Metode <i>Support Vector Machine</i> .....	124
Tabel 147. Confusion Matrix Fold-10 Maskapai Sriwijaya <i>Air</i> Metode <i>Support Vector Machine</i> .....	124
Tabel 148. Nilai TP, FP, FN, TN Masing-masing Fold Metode <i>Support Vector Machine</i> pada Maskapai Sriwijaya <i>Air</i> .....	125
Tabel 149. Pengukuran kinerja <i>Support Vector Machine</i> pada dataset maskapai Sriwijaya <i>Air</i> .....	126
Tabel 150. Hasil Akurasi Masing-Masing Fold Metode <i>Support Vector Machine</i> Pada Dataset Maskapai Penerbangan Indonesia .....	127
Tabel 151. Confusion Matrix Fold-1 Maskapai Penerbangan Indonesia Metode <i>Support Vector Machine</i> .....	128
Tabel 152. Confusion Matrix Fold-2 Maskapai Penerbangan Indonesia Metode <i>Support Vector Machine</i> .....	128
Tabel 153. Confusion Matrix Fold-3 Maskapai Penerbangan Indonesia Metode <i>Support Vector Machine</i> .....	129
Tabel 154. Confusion Matrix Fold-4 Maskapai Penerbangan Indonesia Metode <i>Support Vector Machine</i> .....	129
Tabel 155. Confusion Matrix Fold-5 Maskapai Penerbangan Indonesia Metode <i>Support Vector Machine</i> .....	130
Tabel 156. Confusion Matrix Fold-6 Maskapai Penerbangan Indonesia Metode <i>Support Vector Machine</i> .....	130
Tabel 157. Confusion Matrix Fold-7 Maskapai Penerbangan Indonesia Metode <i>Support Vector Machine</i> .....	131
Tabel 158. Confusion Matrix Fold-8 Maskapai Penerbangan Indonesia Metode <i>Support Vector Machine</i> .....	131

Tabel 159. Confusion Matrix Fold-9 Maskapai Penerbangan Indonesia Metode <i>Support Vector Machine</i> .....	132
Tabel 160. Confusion Matrix Fold-10 Maskapai Penerbangan Indonesia Metode <i>Support Vector Machine</i> .....	132
Tabel 161. Nilai TP, FP, FN, TN Masing-masing Fold Metode <i>Support Vector Machine</i> pada Maskapai Penerbangan Indonesia .....	133
Tabel 162. Pengukuran Kinerja <i>Support Vector Machine</i> pada Dataset Maskapai Penerbangan Indonesia.....	134
Tabel 163. Hasil Akurasi Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Garuda .....	135
Tabel 164. Hasil <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Garuda Kelas Negatif .....	136
Tabel 165. Hasil <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Garuda Kelas Positif.....	137
Tabel 166. Hasil <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Garuda Kelas Negatif .....	138
Tabel 167. Hasil <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Garuda Kelas Positif.....	138
Tabel 168. Hasil <i>F1-Score</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Garuda Kelas Negatif .....	139
Tabel 169. Hasil <i>F1-Score</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Garuda Kelas Positif.....	140
Tabel 170. Hasil Akurasi Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Lion Air.....	141
Tabel 171. Hasil <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Lion Air Kelas Negatif .....	142
Tabel 172. Hasil <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Lion Air Kelas Positif.....	143
Tabel 173. Hasil <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Lion Air Kelas Negatif .....	144
Tabel 174. Hasil <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Lion Air Kelas Positif.....	144

Tabel 175. Hasil F1-Score Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Lion Air</i> Kelas Negatif .....	146
Tabel 176. Hasil F1-Score Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Lion Air</i> Kelas Positif .....	146
Tabel 177. Hasil Akurasi Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Batik Air</i> .....	147
Tabel 178. Hasil <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Batik Air</i> Kelas Negatif .....	148
Tabel 179. Hasil <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Batik Air</i> Kelas Positif.....	149
Tabel 180. Hasil <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Batik Air</i> Kelas Negatif .....	150
Tabel 181. Hasil <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Batik Air</i> Kelas Positif.....	150
Tabel 182. Hasil F1-Score Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Batik Air</i> Kelas Negatif .....	151
Tabel 183. Hasil F1-Score Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Batik Air</i> Kelas Positif.....	152
Tabel 184. Hasil Akurasi Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Sriwijaya Air</i> .....	153
Tabel 185. Hasil <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Sriwijaya Air</i> Kelas Negatif.....	154
Tabel 186. Hasil <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Sriwijaya Air</i> Kelas Positif .....	155
Tabel 187. Hasil <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Sriwijaya Air</i> Kelas Negatif.....	156
Tabel 188. Hasil <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Garuda</i> Kelas Positif.....	156
Tabel 189. Hasil F1-Score Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Sriwijaya Air</i> Kelas Negatif.....	158
Tabel 190. Hasil f1-Score Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai <i>Sriwijaya Air</i> Kelas Positif .....	158

Tabel 191. Hasil Akurasi Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Penerbangan Indonesia .....	160
Tabel 192. Hasil <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Penerbangan Indonesia Kelas Negatif.....	161
Tabel 193. Hasil <i>Precision</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Penerbangan Indonesia Kelas Positif .....	161
Tabel 194. Hasil <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Penerbangan Indonesia Kelas Negatif.....	162
Tabel 195. Hasil <i>Recall</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Penerbangan Indonesia Kelas Positif .....	163
Tabel 196. Hasil <i>F1-Score</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Penerbangan Indonesia Kelas Negatif.....	164
Tabel 197. Hasil <i>F1-Score</i> Metode <i>Naive Bayes</i> dan <i>Support Vector Machine</i> pada Maskapai Penerbangan Indonesia Kelas Positif .....	165

## I. PENDAHULUAN

### 1.1 Latar Belakang

Indonesia, sebagai negara kepulauan, membentang dari Sabang hingga Merauke, terdiri dari ribuan pulau besar dan kecil yang terhubung oleh berbagai selat dan laut. Berdasarkan informasi resmi dari situs Badan Informasi Geospasial, luas total wilayah Indonesia adalah 1.919.440 km<sup>2</sup>. Saat ini, terdapat 13.466 pulau yang terdaftar dan memiliki koordinat. Demi memfasilitasi mobilitas dan konektivitas antar wilayah, keberadaan transportasi sangat diperlukan (Azmarani, 2016).

Transportasi dapat dikelompokkan berdasarkan jenis jalur atau permukaan jalan yang digunakan, alat angkutan yang dipakai, dan tenaga penggerak yang digunakan. Klasifikasi tersebut mencakup transportasi darat, transportasi melalui air, dan transportasi udara. Transportasi udara dianggap sebagai sarana transportasi yang paling modern dan cepat. Transportasi udara menggunakan pesawat udara dalam berbagai jenisnya sebagai alat transportasi, dengan udara atau ruang angkasa sebagai jalurnya. Dalam konteks ini, transportasi udara atau pesawat udara dianggap sebagai jenis transportasi yang paling efektif untuk menghubungkan wilayah-wilayah atau pulau-pulau yang berbeda (Sari & Wibowo, 2019)

Maskapai penerbangan merupakan bisnis khusus dalam mengangkut orang dan kargo melalui udara. Berkat perjanjian kode bersama, yang memungkinkan satu maskapai untuk mengoperasikan penerbangan yang sama dengan menggunakan kode yang sama, maskapai dapat memberikan berbagai layanan kepada pelanggan mereka. Sertifikat atau lisensi operasi udara yang diberikan oleh organisasi penerbangan pemerintah biasanya diakui oleh maskapai penerbangan (Airin et al., 2022).

Industri penerbangan merupakan sektor yang penuh inovasi, mampu memberikan dukungan bagi nilai ekonomi dan sosial. memiliki kemampuan untuk menghubungkan orang, negara, dan budaya, memberikan akses ke pasar global, memfasilitasi perdagangan dan pariwisata, serta memperkuat hubungan antara negara maju dan negara berkembang. Selain itu, industri ini menyediakan jaringan transportasi global yang cepat, yang sangat penting untuk kebutuhan bisnis dan pariwisata global, serta berperan kunci dalam memfasilitasi pertumbuhan ekonomi, terutama di negara-negara berkembang. Pertumbuhan ekonomi ini diharapkan dapat meningkatkan standar hidup dan mengurangi tingkat kemiskinan (Prissia & Daryanto, 2019).

Kondisi persaingan di industri penerbangan nasional terus meningkat dan sangat ketat, terutama dengan penambahan kapasitas operator penerbangan berbiaya murah baik untuk rute domestik maupun internasional. Hal ini juga diiringi dengan penerapan kebijakan ruang udara terbuka ASEAN secara bertahap, yang semakin memperumit dinamika persaingan di industri ini (Prissia & Daryanto, 2019). Tingkat kepuasan pelanggan mencerminkan sejauh mana produk, layanan, atau pengalaman yang diberikan oleh suatu perusahaan dapat memenuhi harapan pelanggan. Kepuasan pelanggan melibatkan aspek psikologis, baik secara umum maupun spesifik terkait dengan pengalaman pelanggan terhadap produk dan layanan perusahaan. Hal ini berarti bahwa umpan balik pelanggan dapat dipengaruhi oleh sentimen dan emosi yang dirasakan oleh pelanggan (Nahumury et al., 2022).

Penting untuk disadari bahwa kepuasan pelanggan memegang peranan kunci dalam menjaga kelangsungan bisnis dalam jangka waktu yang panjang, seiring dengan upaya membangun kualitas dan layanan perusahaan. Bisnis yang berhasil mencapai tingkat kepuasan pelanggan yang tinggi cenderung memiliki kualitas layanan yang unggul. Bagi perusahaan, terutama maskapai penerbangan, penggunaan kuesioner, baik yang dilakukan secara tradisional melalui brosur maupun secara online, sering menjadi metode yang umum digunakan untuk memperoleh umpan balik pelanggan. Proses pengumpulan data kepuasan pelanggan atau analisis sentimen secara



tradisional dianggap mudah, tetapi kurang efektif karena banyak responden yang memberikan jawaban yang tidak tepat dan tidak relevan, yang dapat membuat data menjadi tidak valid dan harus dibuang saat pembersihan data (Nahumury et al., 2022).

Di sisi lain, media sosial *X* telah menjadi media yang digunakan oleh maskapai penerbangan sebagai sumber untuk mengumpulkan sentimen pelanggan terkait produk dan layanan. Hal ini dikarenakan sifatnya yang instan dan dapat diandalkan, membuatnya menjadi *platform* yang efisien untuk memahami pandangan dan perasaan pelanggan (Nahumury et al., 2022).

Data yang akan dimanfaatkan dalam penelitian ini berasal dari *tweet* yang terkait dengan layanan dari maskapai penerbangan di Indonesia. Maskapai ini mendapatkan sentimen positif atau apresiasi karena harganya yang terjangkau, menawarkan rute yang meliputi sebagian besar wilayah Indonesia, dan telah memperoleh sertifikasi *Operational Safety Audit* (IOSA). Namun, terdapat juga sentimen negatif dari pelanggan terkait dengan keterlambatan penerbangan, penghapusan bagasi gratis, dan kecelakaan pesawat (Nahumury et al., 2022).

Dalam *tweet* yang diposting oleh pelanggan maskapai penerbangan Indonesia, data tekstual atau pengetahuan dapat ditemukan dan diekstraksi menggunakan teknik text mining. Data tersebut kemudian akan dianalisis dengan menggunakan alat-alat yang telah diteliti sebelumnya, sehingga dapat terungkap ide dan sentimen pelanggan terkait dengan tingkat kepuasan mereka. Penelitian ini bertujuan untuk evaluasi kinerja maskapai penerbangan di Indonesia (Nahumury et al., 2022).

Media sosial *X* merupakan *platform* media sosial yang berfungsi untuk interaksi antar pengguna. Konsep utama *media sosial X* adalah menyebarkan informasi dalam bentuk pesan singkat, padat, dan *real-time*, dengan kalimat yang terbatas pada 140 karakter atau kurang. Pesan-pesan ini dapat diakses oleh pembaca di seluruh dunia. *media sosial X* juga berfungsi sebagai sarana efektif untuk menyebarkan informasi kepada semua orang, tanpa memandang apakah pengguna

tersebut dikenal atau tidak dikenal. Tujuan utama penggunaan *media sosial X* adalah memberitahukan keberadaan dan berbagi informasi dengan cepat (Muludi et al., 2021).

Media sosial *X* memberikan kemudahan bagi pengguna untuk berinteraksi dan berkoneksi dengan pengguna lain. Media sosial *X* juga dilengkapi dengan Fitur *Top Trending*, yang mempermudah pengguna untuk melihat *tweet-tweet* yang sedang populer dan banyak dibicarakan oleh pengguna *media sosial X* (Muludi et al., 2021).

Pemrosesan bahasa alami, analisis teks, komputasi linguistik, dan biometrik digunakan dalam analisis sentimen untuk melakukan analisis, ekstraksi, pengukuran, dan pendeteksian secara sistematis terhadap keadaan afektif dan informasi pribadi. Secara tradisional, analisis sentimen berfokus pada polaritas opini, menilai apakah suatu opini memiliki sifat positif, netral, atau negatif terhadap suatu hal atau layanan. Produk yang sering diulas di internet menjadi fokus utama dari analisis sentimen (Airin et al., 2022).

Analisis sentimen adalah teknik yang digunakan untuk mengekstrak informasi dari data teks dengan tujuan mendapatkan pemahaman mengenai sentimen yang dapat berupa positif, netral, atau negatif. Analisis sentimen melibatkan penilaian atau pendapat pribadi yang diberikan oleh pengguna internet melalui media sosial. Dalam upaya untuk menjelajahi dan memahami pandangan masyarakat, analisis sentimen menjadi alat yang efektif dalam mengidentifikasi dan memahami opini serta sikap yang muncul dalam percakapan di media sosial. Dengan menerapkan algoritma *Naive Bayes* dan *Support Vector Machine*, kita dapat memperoleh pemahaman yang lebih komprehensif tentang bagaimana masyarakat menanggapi isu tersebut (Sari & Wibowo, 2019)

*Naive Bayes Classifier* merupakan sebuah pengklasifikasi statistik *Bayesian* yang sederhana. Judea Pearl mengemukakan *Bayesian Network* (BN) sebagai model grafis yang mengkodekan distribusi probabilitas gabungan dari sekumpulan

variabel acak diskrit (Ruz et al., 2020). Istilah "*Naive*" digunakan karena pendekatan ini mengasumsikan bahwa semua variabel memberikan kontribusi terhadap klasifikasi dan saling berkorelasi. Metode ini juga sering disebut dengan sebutan *Idiot's Bayes*, *Simple Bayes*, dan *Independence Bayes*. Pengklasifikasi *Naive Bayes* dapat memprediksi probabilitas keanggotaan kelas, seperti probabilitas bahwa suatu data termasuk dalam label kelas tertentu. Pendekatan *Naive Bayes* menganggap bahwa keberadaan (atau ketiadaan) fitur (atribut) tertentu dari suatu kelas tidak memiliki keterkaitan dengan keberadaan (atau ketiadaan) fitur lainnya ketika variabel kelas diberikan (Jadhav & Channe, 2016).

*Support Vector Machines* (SVM) merupakan metode yang baru-baru ini dikembangkan dalam kerangka teori pembelajaran statistik oleh Vapnik pada tahun 1998. Metode ini telah berhasil diaplikasikan dalam berbagai konteks, mulai dari prediksi deret waktu hingga pengenalan wajah, bahkan dalam pemrosesan data biologis untuk diagnosis medis. Keberhasilan eksperimental dan dasar teoritis SVM mendorong penelitian lebih lanjut untuk memahami karakteristiknya dan memperluas penerapannya ke berbagai bidang (Evgeniou & Pontil, 2001)

*Support Vector Machine* (SVM) juga merupakan algoritma yang memiliki dasar teoritis yang kuat dan telah berkembang sejak tahun 60-an melalui teori pembelajaran statistik (Pravina et al., 2019). SVM termasuk dalam kelas algoritma yang digunakan untuk pengenalan pola dan dikenal sebagai alat pembelajaran klasifikasi yang efektif. SVM merupakan perluasan dari model nonlinier dari algoritma umum yang dikembangkan oleh Vladimir Vapnik, seorang ahli dalam teori pembelajaran statistik. Algoritma SVM didasarkan pada teori pembelajaran statistik dan dimensi Vapnik-Chervonenkis (VC) yang diperkenalkan oleh Vladimir Vapnik dan Alexey Chervonenkis. *Support Vector Machines* (SVM) adalah sekelompok metode pembelajaran terawasi yang dapat diterapkan pada tugas klasifikasi atau regresi (Sahayak et al., 2015).

Dalam penelitian ini, akan menganalisis data hasil observasi pada media sosial *X* terkait maskapai penerbangan Indonesia menggunakan dua metode yaitu *Naive*

*Bayes* dan *Support Vector Machine* (SVM). Selanjutnya, akan membandingkan kinerja kedua metode tersebut dalam penelitian ini. Berdasarkan uraian yang telah dijelaskan, judul penelitian ini adalah “Perbandingan Kinerja Algoritma *Naive Bayes* dan *Support Vector Machine* Pada Analisis Sentimen Analisis Sentimen di Media Sosial *media sosial X* terhadap Maskapai Penerbangan Indonesia”.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, penelitian ini berfokus pada dua rumusan masalah, yaitu :

1. Bagaimana sentimen masyarakat terhadap maskapai penerbangan Indonesia yang diekspresikan melalui media sosial *X*?
2. Bagaimana perbandingan kinerja algoritma *Naive Bayes* dan *Support Vector Machine* dalam melakukan analisis sentimen di media sosial *X* terhadap maskapai penerbangan Indonesia?

## 1.3 Batasan Masalah

Adapun batasan masalah pada penelitian ini adalah sebagai berikut.

1. Data *tweet* yang digunakan adalah *tweet* dengan kata kunci yang berhubungan dengan maskapai penerbangan Indonesia.
2. Data *tweet* yang digunakan berbahasa Indonesia dan fokus pada data *tweet* maskapai Garuda, Lion *Air*, Sriwijaya *Air*, dan Batik *Air* sebagai maskapai penerbangan Indonesia.

## 1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut :

1. Melakukan analisis sentimen terhadap maskapai penerbangan Indonesia berdasarkan data yang didapatkan di media sosial *media sosial X*.
2. Melakukan perbandingan hasil akurasi, presisi, *Recall*, dan *F1-Score* yang menggunakan algoritma *Naive Bayes* dan *Support Vector Machine* (SVM) dalam analisis sentiment maskapai penerbangan Indonesia.

## 1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut :

1. Mengetahui performa dari algoritma *Naive Bayes* dan *Support Vector Machine* (SVM) dalam melakukan analisis sentimen terhadap kinerja maskapai penerbangan Indonesia.
2. Mendapatkan informasi mengenai opini masyarakat terhadap kualitas kinerja maskapai penerbangan Indonesia untuk dapat digunakan sebagai bahan pengembangan maskapai penerbangan Indonesia.

## II. TINJAUAN PUSTAKA

### 2.1 Maskapai Penerbangan

Maskapai penerbangan merupakan penyedia layanan transportasi yang menjadi kebutuhan utama bagi masyarakat dalam melakukan perpindahan dari suatu lokasi ke lokasi lain, baik itu antar kota maupun antar negara (Pravina et al. 2019). Menurut penelitian yang dilakukan oleh Sari & Wibowo (2014), jasa penerbangan merupakan salah satu jenis layanan yang memenuhi kebutuhan manusia untuk berpindah dari satu lokasi ke lokasi lain yang jaraknya cukup jauh, dengan waktu tempuh yang singkat. Menurut informasi dari Departemen Perhubungan (2015), data statistik mengenai lalu lintas angkutan udara di Kementerian Perhubungan Republik Indonesia menunjukkan peningkatan jumlah penumpang setiap tahunnya, baik untuk rute domestik maupun internasional.

Berdasarkan data dari Kementrian Perhubungan dalam laporan *Civil Aircraft Register* yang terdaftar per 31 Desember 2022 adalah sebagai berikut.

Tabel 1. Jumlah Pesawat Milik Maskapai Penerbangan Indonesia

No	Maskapai Penerbangan Indonesia	Jumlah Pesawat Milik Maskapai Penerbangan Indonesia (Unit)
1	Garuda	87
2	Batik Air	72
3	Lion Air	112
4	Sriwijaya Air	8
5	Citilink	62

Berdasarkan data pada Tabel 1, *Lion Air* merupakan maskapai penerbangan Indonesia dengan jumlah unit pesawat terbang terbanyak. Diurutan berikutnya ada Garuda, Batik Air, Citilink, dan Sriwijaya Air. Sriwijaya Air tercatat sebagai maskapai dengan jumlah armada pesawat paling sedikit di antara maskapai penerbangan Indonesia lainnya.

Media sosial *X* memiliki banyak pengguna di Indonesia, yang sebagian besar memposting dan menulis *tweet* untuk mengungkapkan pendapat mereka dalam bahasa Indonesia. Untuk memperoleh data dari penumpang mengenai opini mereka terhadap maskapai penerbangan Indonesia (Airin et al., 2022). Penulis mengumpulkan data *tweets* tersebut dengan menggunakan kata kunci yang berhubungan dengan maskapai penerbangan Indonesia. Data yang diperoleh mengacu pada masing-masing maskapai penerbangan Garuda, *Lion Air*, *Batik Air*, *Sriwijaya Air*.

## 2.2 *Text Mining*

Text Mining adalah pendekatan umum untuk menganalisis data dan informasi dari sumber teks yang belum pernah dilihat sebelumnya, dengan tujuan menemukan pola-pola tertentu untuk tujuan tertentu. Dalam implementasinya, text mining sering digunakan untuk menyelidiki data dan berfungsi sebagai alat dalam menggali informasi dari berbagai teks. Sebagai langkah awal, penting untuk melakukan proses preprocessing data yang terdiri dari beberapa metode untuk membersihkan dan menstandarisasi data teks yang akan diolah (Husada & Paramita, 2021).

Menurut Husada & Paramita (2021), beberapa metode dalam preprocessing data teks melibatkan tahapan seperti *cleaning*, *Case Folding*, tokenizing, stopwords removal, dan *stemming*. Tahapan-tahapan ini membantu menyederhanakan dan membersihkan data teks agar lebih mudah diolah dalam analisis selanjutnya. Text mining bertujuan untuk memperoleh informasi yang berkualitas tinggi dari teks. Proses ini umumnya dilakukan dengan maksud untuk mengekstraksi pengetahuan dan informasi yang terkandung dalam pola-pola teks dokumen.

### 2.3 Analisis Sentimen

Analisis sentimen merupakan bagian dari Natural Language Processing (NLP) dan merupakan suatu proses yang digunakan untuk mengidentifikasi opini atau pandangan (sentimen) dalam dataset berupa teks terhadap suatu isu atau kejadian, yang dapat bersifat positif, negatif, atau netral. Analisis sentimen termasuk dalam bidang penelitian yang populer dan diakui memberikan keuntungan dalam berbagai aspek. Secara umum, aplikasi analisis sentimen mencakup prediksi harga saham, isu politik, kepuasan terhadap produk atau layanan, analisis reputasi, dan sebagainya. Salah satu bentuk analisis sentimen yang umum dilakukan adalah mengenai opini terhadap suatu isu yang ditemukan di platform *media sosial X* (Fikri et al. 2020).

### 2.4 Media Sosial X

Media sosial *X*, sebagai salah satu platform media sosial yang sangat populer, memiliki peran yang signifikan sebagai tempat berkomunikasi dalam masyarakat. Dengan menggunakan media sosial *X*, individu di seluruh dunia dapat terhubung dengan keluarga, teman, dan kenalan mereka melalui perangkat komputer atau ponsel mereka. Media sosial *X* menyediakan layanan pembuatan pesan status yang disebut "*tweets*," yang dapat dibaca oleh pengguna media sosial *X* lainnya. Dengan demikian, *media sosial X* menjadi salah satu platform yang mengumpulkan data opini dari masyarakat di seluruh dunia (Fikri et al. 2020).

*Tweets*, yang muncul sebagai hasil dari ekspresi opini dan komentar, menjadi sumber daya yang berguna untuk menganalisis sentimen masyarakat umum terhadap suatu instansi atau individu. Hal ini disebabkan oleh fakta bahwa *tweets* tersebut mengandung sentimen yang dapat menjadi indikator pandangan umum masyarakat yang dapat dijadikan sebagai bahan evaluasi (Fikri et al. 2020).

Menurut laporan dari *We Are Social* pada <https://databoks.katadata.co.id/> (diakses pada 08:17 WIB 06 Maret 2024) terdapat 10 media sosial yang paling banyak digunakan di Indonesia per Januari 2024.



Tabel 2. 10 Aplikasi Dengan Pengguna Terbanyak di Indonesia

No	Media Sosial	Pengguna (%)
1	<i>Whatsapp</i>	90,9
2	<i>Instagram</i>	85,3
3	<i>Facebook</i>	81,6
4	<i>Tiktok</i>	73,5
5	<i>Telegram</i>	61,3
6	<i>media sosial X</i>	57,5
7	<i>Facebook Messenger</i>	47,9
8	<i>Pinterest</i>	34,2
9	<i>Kuaishou</i>	32,4
10	<i>Linkedin</i>	25

Secara keseluruhan tercatat 139 juta identitas pengguna media sosial di Indonesia pada Januari 2024. Pada Tabel 2, terlihat *whatsapp* menjadi aplikasi dengan pengguna terbanyak per Januari 2024 dan media sosial *X* berada pada urutan ke 6 sebagai aplikasi dengan pengguna terbanyak di Indonesia.

## 2.5 K-Fold cross validation

*K-fold cross-validation* adalah salah satu teknik yang digunakan untuk membagi data menjadi data *training* dan *testing*. Teknik ini banyak digunakan oleh para peneliti karena terbukti dapat mengurangi bias yang muncul akibat pengambilan sampel. *K-fold cross-validation* secara berulang-ulang membagi data menjadi data *training* dan data *testing*, sehingga setiap data memiliki kesempatan untuk menjadi data *testing*. *K* adalah jumlah pembagian yang digunakan untuk membagi data menjadi data *training* dan data *testing*. Berikut adalah ilustrasi dari penerapan *K-Fold cross validation* dengan nilai *K* adalah 5 dapat dilihat pada Tabel 3 (Husada & Paramita, 2021).

Tabel 3. Ilustrasi pembagian data menggunakan *k-fold cross validation*

<i>Fold-1</i>	<i>Fold-2</i>	<i>Fold-3</i>	<i>Fold-4</i>	<i>Fold-5</i>
<i>Test</i>	<i>Train</i>	<i>Train</i>	<i>Train</i>	<i>Train</i>
<i>Train</i>	<i>Test</i>	<i>Train</i>	<i>Train</i>	<i>Train</i>
<i>Train</i>	<i>Train</i>	<i>Test</i>	<i>Train</i>	<i>Train</i>
<i>Train</i>	<i>Train</i>	<i>Train</i>	<i>Test</i>	<i>Train</i>
<i>Train</i>	<i>Train</i>	<i>Train</i>	<i>Train</i>	<i>Test</i>

## 2.6 *Naive Bayes*

Algoritma *Naive Bayes* adalah metode yang umum digunakan dalam klasifikasi teks. Konsep dasarnya melibatkan kombinasi probabilitas kata-kata dan kategori untuk memperkirakan probabilitas kategori suatu dokumen. *Naive Bayes* merupakan algoritma klasifikasi probabilistik yang paling sederhana. Meskipun *Naive Bayes* dikenal sebagai metode yang sangat sederhana dan efisien, algoritma ini sangat sensitif terhadap pemilihan fitur seleksi. Oleh karena itu, pemilihan fitur yang tepat menjadi sangat penting dalam implementasinya (Nurdina & Puspita, 2023).

*Naive Bayes* merupakan salah satu algoritma yang sangat efisien dan efektif dalam konteks pembelajaran mesin dan analisis data, menggunakan probabilitas dan perhitungan statistik yang dikembangkan oleh ilmuwan Inggris, Thomas Bayes. Dalam implementasinya pada dataset penelitian, algoritma ini dijadikan dasar

sistem yang menerapkan *Naive Bayes*. Fitur utama dari *Naive Bayes Classifier* adalah asumsi yang sangat kuat (naïf) mengenai independensi setiap kondisi atau kejadian. Meskipun sederhana, *Naive Bayes* terbukti akurat dan cepat ketika diterapkan pada *database* yang besar. Keuntungan menggunakan *Naive Bayes* terletak pada kebutuhan informasi pelatihan yang minim, hanya memerlukan sedikit data untuk menentukan parameter mean dan varians yang diperlukan untuk proses klasifikasi (Nurdina & Puspita, 2023).

Perhitungan dalam teori *Bayes* menggunakan persamaan berikut (Fikri et al., 2020):

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (1)$$

Keterangan :

X = Data dengan kelas yang belum diketahui

C = Hipotesis data X merupakan kelas spesifik

$P(C|X)$  = Probabilitas hipotesis C berdasarkan kondisi X (*Posterior Probability*)

$P(C)$  = Probabilitas hipotesis C (*Prior Probability*)

$P(X|C)$  = Probabilitas X berdasarkan kondisi pada hipotesis C (*Likelihood*)

$P(X)$  = Probabilitas X (*Predictor prior probability*)

Pada persamaan 1 dapat dijabarkan menjadi beberapa tahapan persamaan, sebagai berikut (Santoso et al., 2022):

$$P(X) = \frac{d_x}{d} \quad (2)$$

Keterangan :

X = Hipotesa data merupakan suatu kelas

$P(X)$  = *Prior probability*

$d$  = Jumlah seluruh dokumen

$d_x$  = Jumlah seluruh data dengan kategori kelas X

$$P(\text{Term } H_n|X) = \frac{(H_n|X)+1}{(c)+|V|} \quad (3)$$

$P(\text{Term } H_n|X)$  = *Conditional Probability* kata ke-n dengan kelas c

$(H_n|X)$  = Frekuensi kemunculan kata pada kelas c

(c) = Bobot (*TF-IDF*) dari kata kelas c

|V| = Nilai *TF* dari seluruh kata yang terdapat pada dokumen

Pada Persamaan 2, penambahan angka 1 pada pembilang berfungsi untuk mengatasi masalah ketika data baru tidak mengandung istilah yang muncul pada data pelatihan, sehingga nilai pada pembilang tidak menjadi nol.

$$P(\text{Term } H|X) = P(X) \times P(\text{Term } H_1|X) \times P(\text{Term } H_n|X) \quad (4)$$

Keterangan :

$P(\text{Term } H|X)$  = *Posterior probability*

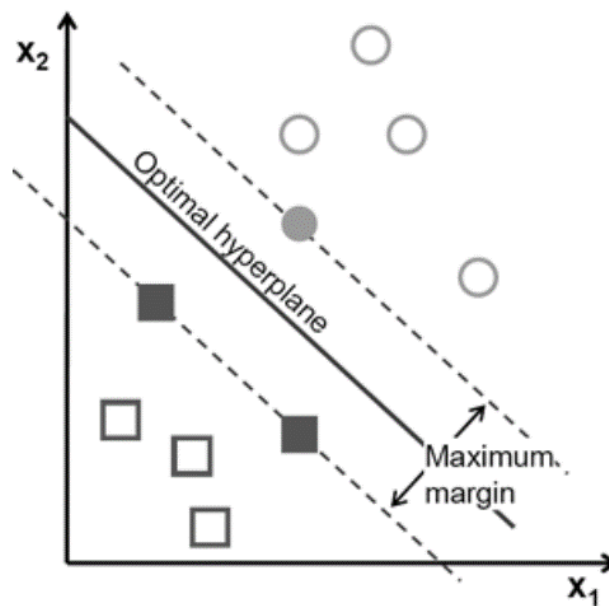
$P(X)$  = *Prior probability*

$P(\text{Term } H_n|X)$  = *Conditional probability* ke-n dengan kategori kelas X

## 2.7 *Support Vector Machine*

*Support Vector Machine* (SVM) adalah algoritma *Machine Learning* yang menggunakan fungsi *hyperplane* untuk memisahkan data ke dalam daerah-daerah kelas. *Hyperplane* sendiri berfungsi sebagai pemisah antara kelas-kelas yang ada. Saat melakukan prediksi kelas dari suatu data, SVM akan memberikan label berdasarkan daerah kelas di mana data tersebut berada. Algoritma SVM sering digunakan pada dataset yang besar, khususnya yang diambil dari sumber online, dan menjadi populer terutama dalam konteks klasifikasi teks (Fikri et al. 2020).

*Support Vector Machine* (SVM) adalah salah satu dari berbagai algoritma yang digunakan untuk klasifikasi dan termasuk dalam kategori *supervised learning*. Prinsip kerja SVM melibatkan pencarian *hyperplane* atau garis pembatas optimal yang berfungsi untuk memisahkan dua kelas (Husada & Paramita, 2021).



Gambar 1. Ilustrasi *Support Vector Machine*.

Sumber: (Husada & Paramita, 2021)

Pada Gambar 1, terlihat struktur *Support Vector Machine* (SVM) terdiri dari dua kelas data, yakni kelas +1 dan kelas -1. *Hyperplane* berfungsi sebagai pemisah antara kedua kelas tersebut. Data yang berada paling dekat dengan garis *hyperplane* dibatasi oleh margin, dan data yang berada pada margin tersebut dikenal sebagai *support vector* (Husada & Paramita, 2021).

Dalam suatu dataset, variabel  $x_i$  mewakili data, sedangkan variabel  $y_i$  mewakili kelas dari data tersebut. Metode SVM membagi dataset menjadi dua kelas. Satu kelas dipisahkan oleh *hyperplane* dan diberi nilai 1, sedangkan kelas lainnya diberi nilai -1. Penentuan prediksi apakah dataset berada pada kelas +1 atau kelas -1 dapat dilihat pada persamaan berikut.

$$X_i \cdot W + b \geq 1 \text{ untuk } Y_i = 1 \quad (2)$$

$$X_i \cdot W + b \leq -1 \text{ untuk } Y_i = -1 \quad (3)$$

$$b = -\frac{1}{2}(w \cdot x^+ + w \cdot x^-) \quad (4)$$

Keterangan :

$X_i$  = data ke-i

$W$  = nilai bobot *support vector*

$b$  = nilai bias

$Y_i$  = kelas data ke- $i$

$w \cdot x^+$  = nilai bobot untuk kelas data positif

$w \cdot x^-$  = nilai bobot untuk kelas data negatif

## 2.8 Confusion Matrix

*Confusion matrix* adalah alat yang digunakan untuk mengevaluasi kinerja model klasifikasi. Matriks ini memungkinkan perbandingan antara hasil kelas yang diprediksi dan label kelas yang sebenarnya. Perbandingan ini sangat penting untuk menghitung metrik seperti akurasi, *Precision*, *Recall*, dan *f1-score*. Berikut adalah ilustrasi *confusion matrix* dapat dilihat pada Tabel 4 (Husada & Paramita, 2021).

Tabel 4. Penerapan *Confusion Matrix* untuk evaluasi kinerja model

<i>Confusion Matrix</i>		<i>Prediction Class</i>	
		Negatif	Positif
<i>Actual Class</i>	Negatif	TN	FP
	Positif	FN	TP

Keterangan :

TN (*True Negative*) : Data negatif yang terklasifikasi secara benar.

FP (*False Positive*) : Data negatif yang terklasifikasi menjadi positif.

FN (*False Negative*) : Data positif yang terklasifikasi menjadi negatif.

TP (*True Positive*) : Data positif yang terklasifikasi secara benar.

## 2.9 Python

*Python* merupakan bahasa yang sangat cocok untuk keperluan pembelajaran dan pengembangan aplikasi dunia nyata. *Python* adalah bahasa pemrograman tingkat tinggi yang berorientasi objek dan memiliki kekuatan yang tinggi. Bahasa ini

dirancang oleh Guido van Rossum. Beberapa *library* yang dapat digunakan *python* adalah sebagai berikut (Srinath, 2017).

### **1. *Pandas***

*Pandas* adalah sebuah pustaka *Python* yang digunakan untuk memanipulasi dan menganalisis data yang terstruktur. Untuk menggunakan fungsi-fungsi dari *Pandas*, kita dapat mengimpornya dengan perintah "*import pandas as pd*".

### **2. *Matplotlib***

*Matplotlib* adalah sebuah pustaka *Python* yang digunakan untuk memvisualisasikan hasil analisis data dalam bentuk grafik dan chart.

### **3. *Wordcloud***

*Wordcloud* merupakan pustaka *Python* yang digunakan untuk menampilkan hasil analisis dalam bentuk gambar *wordcloud*. Di dalam pustaka ini, terdapat fitur *stopwords* yang berguna untuk mengeliminasi kata-kata umum yang sering muncul, namun tidak memiliki signifikansi dalam hasil analisis data

## 2.8. Penelitian Terdahulu

Tabel 5. Penelitian Terdahulu

No	Judul Penelitian	Data	Metode	Hasil
1	Analisis Sentimen Tentang Opini Masyarakat Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritma <i>Support Vector Machine</i> (Husada, H., & Paramita, A.,2021).	Data yang digunakan merupakan data yang diambil dari <a href="https://www.figure-eight.com/dota-for-everyone/">https://www.figure-eight.com/dota-for-everyone/</a> .	<i>Support Vector Machine</i>	Akurasi yang didapat sebesar 84,37 %, nilai presisi, <i>Recall</i> , dan <i>f-Score</i> masing-masing sebesar 84,33%, 84,67%, dan 84,50%.
2	<i>The Best Malaysian Airline Companies Visualization through Bilingual Twitter Sentiment Analysis: A Machine Learning Classification</i> (Airin et al., 2022)	Data yang digunakan merupakan data yang diambil dari <i>Kaggle</i> dan <i>Github</i>	<i>Naive Bayes</i>	Akurasi yang didapat sebesar 93%.



Lanjutan Tabel 5. Penelitian Terdahulu

<b>No</b>	<b>Judul Penelitian</b>	<b>Data</b>	<b>Metode</b>	<b>Hasil</b>
3	<i>Drug Sentiment Analysis using Machine Learning Classifiers</i> (Uddin et al., 2022)	Data yang digunakan dalam penelitian ini adalah data yang diambil UCI repository	<i>Naïve Bayes, Support Vector Machine, dan Random Forest</i>	Akurasi yang didapatkan dari <i>Drug Sentiment Analysis</i> menggunakan <i>Naïve Bayes, Support Vector Machine, dan Random Forest</i> masing-masing sebesar 88,61%, 88,42%, 94,11%
4	<i>Naive Bayes and KNN for Airline Passenger Satisfaction Classification: Comparative Analysis</i> (Nurdina, A., & Puspita, A., 2023)	Data yang digunakan dalam penelitian ini adalah data yang diperoleh dari <i>Kaggle</i>	<i>Naive Bayes dan K-nearest Neighbour</i>	Hasil akurasi yang didapat menggunakan metode <i>Naïve Bayes dan K-nearest Neighbour</i> masing-masing sebesar 84,48% dan 65,38%

### III. METODE PENELITIAN

#### 3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan di Jurusan Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung. Berlokasi di Jalan Soemantri Brojonegoro No. 1, Gedung Meneng, Bandar Lampung, penelitian ini akan dilaksanakan pada Maret 2024 hingga bulan September 2024.

#### 3.2 Perangkat Penelitian

Perangkat yang digunakan untuk penelitian ini adalah sebagai berikut.

1. Perangkat Keras

Laptop Lenovo V330 dengan SSD 256 GB, RAM 20 GB, dan *Processor* AMD Ryzen 3 2200U

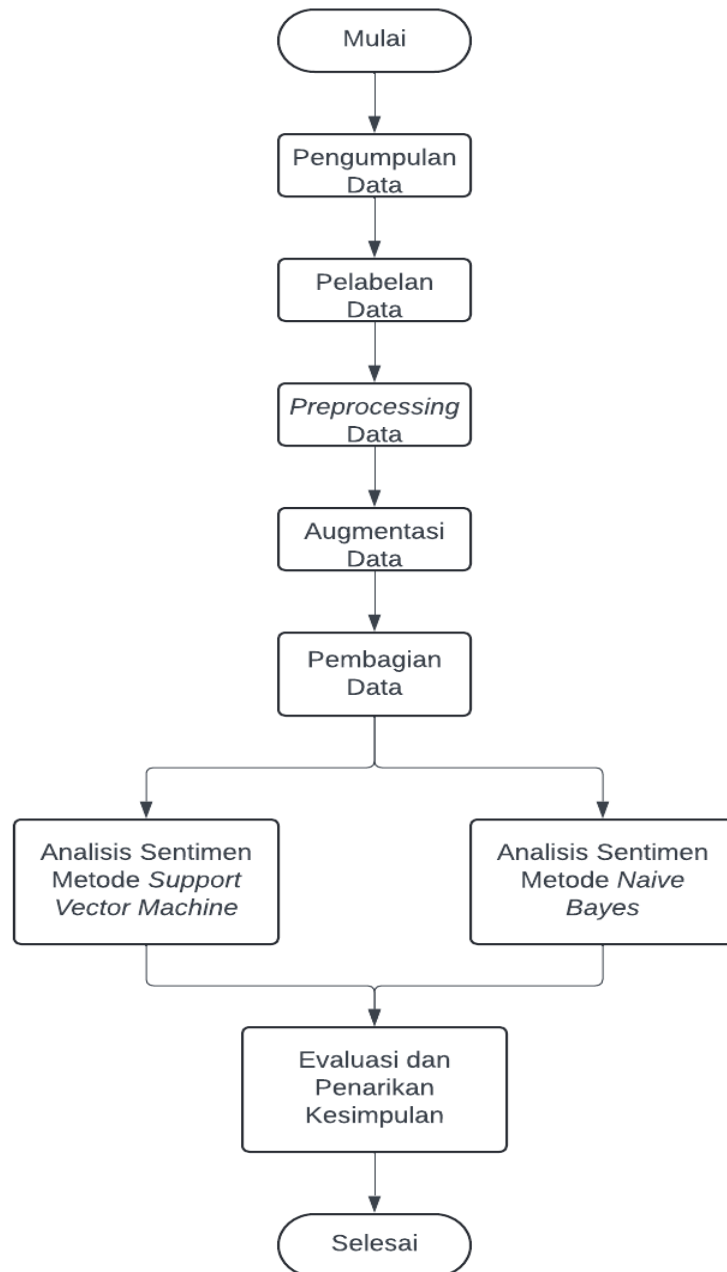
2. Perangkat Lunak

Perangkat lunak yang digunakan dalam penelitian adalah sebagai berikut.

- *Windows* 10 64-Bit sebagai sistem operasi laptop
- *Web Browser Google Chrome, Python, dan Microsoft Excel* sebagai perangkat lunak pendukung dalam proses analisis sentimen

### 3.3 Tahapan Penelitian

Tahapan penelitian ini dapat dilihat pada Gambar 2.



Gambar 2. Tahapan Penelitian.

Gambar 2, menunjukkan proses yang akan dilalui dalam penelitian ini. Berikut adalah penjabaran dari masing-masing tahapan penelitian.

## **I. Pengumpulan Data**

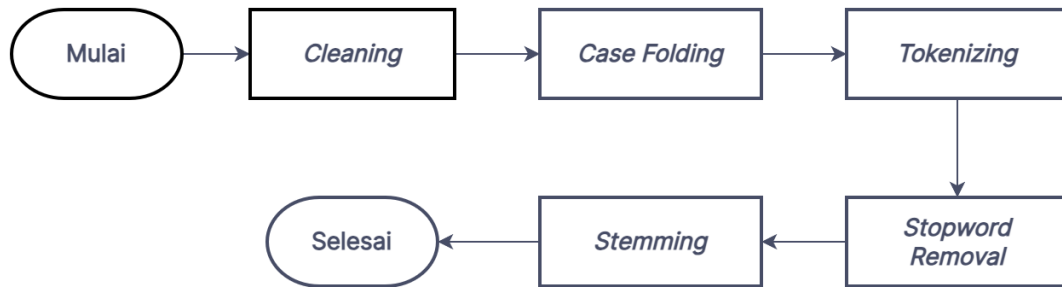
Tahapan pertama dalam penelitian ini adalah pengumpulan data untuk analisis sentiment terhadap maskapai penerbangan Indonesia. Dataset yang digunakan untuk penelitian ini diambil dari media sosial *X*. Kata kunci yang digunakan untuk pengambilan data adalah kata kunci yang berhubungan dengan maskapai penerbangan Indonesia, seperti “Maskapai penerbangan Indonesia”, “Garuda”, “Lion Air”, “Batik Air”, “Sriwijaya Air”. Data *tweet* yang terkumpul kemudian disimpan dalam format csv. Data yang diambil adalah data *tweet* yang dimulai dari bulan Januari 2023 sampai dengan bulan Desember 2023.

## **II. Pelabelan Data**

Pada tahap ini, dataset yang telah dikumpulkan akan diberi label sentimen positif atau sentimen negatif. Pemberian nilai sentimen positif dan negatif dilakukan secara manual pada data *tweet* yang telah dikumpulkan. Pemberian nilai sentimen positif didasari pada kalimat yang memiliki unsur positif, bersifat membangun, dan tidak memiliki unsur kebencian. Pemberian nilai sentimen negatif didasari pada kalimat yang memiliki unsur negatif, bersifat menjatuhkan, dan memiliki unsur kebencian.

## **III. *Preprocessing* Data**

*Preprocessing* data adalah proses pembersihan atau penyesuaian data *tweet* agar dapat diolah lebih lanjut. Data yang dikumpulkan mungkin masih dalam bentuk yang tidak terstruktur atau berantakan. Proses ini melibatkan penghapusan informasi yang tidak diperlukan, standarisasi format, dan langkah-langkah lainnya untuk memastikan bahwa data tersebut siap untuk analisis lebih lanjut. Berikut adalah alur *preprocessing* yang dapat dilihat pada Gambar 3.



Gambar 3. Alur preprocessing data.

Berdasarkan Gambar 3, berikut adalah penjelasan alur dari *preprocessing* data.

### 1. *Cleaning*

*Cleaning* merupakan proses untuk menghilangkan *noise* atau gangguan pada data. Pada tahap ini, semua data teks penelitian akan melalui tahap pembersihan data seperti menghapus *username*, URL, tanda *hashtag* (#), tanda *mention* (@), angka, tanda baca, karakter HTML, dan simbol lainnya. Tujuan dilakukannya *cleaning* adalah untuk menghilangkan elemen-elemen yang tidak relevan atau mengganggu dalam teks, sehingga hanya informasi teks yang penting yang tetap terjaga. Selain itu, atribut-atribut yang tidak memberikan kontribusi signifikan juga akan dihilangkan.

### 2. *Case Folding*

*Case Folding* merupakan proses standarisasi yang mentransformasikan semua komponen dalam teks menjadi huruf kecil (*lowercase*). Selain itu, dalam proses ini, karakter-karakter selain huruf yang dianggap tidak memiliki dampak signifikan pada pengolahan teks, seperti angka dan tanda baca, juga akan dihapus. Tujuan dari *Case Folding* adalah untuk menormalkan kata atau kalimat dalam data. Fungsinya adalah untuk memperbaiki struktur kata dan kosa kata dalam suatu kalimat.

### 3. *Tokenizing*

Tokenisasi adalah proses di mana kata-kata yang tersusun bersama dalam sebuah kalimat dipisahkan menjadi unit-unit kata tunggal atau yang dikenal sebagai token. Dalam proses tokenisasi, dilakukan pemotongan string input berdasarkan setiap kata yang membentuknya. Secara umum, tokenisasi melibatkan pemecahan karakter-karakter dalam suatu teks ke dalam unit kata, dengan memperhatikan karakter-karakter tertentu yang dapat dianggap sebagai pemisah

kata. Pada umumnya, setiap kata diidentifikasi atau dipisahkan dari kata lainnya oleh karakter spasi. Oleh karena itu, proses tokenisasi bergantung pada karakter spasi dalam dokumen untuk memisahkan kata-kata.

#### **4. *Stopword Removal***

*Stopword Removal* adalah proses penghilangan kata-kata yang tidak diperlukan dalam data, bertujuan untuk meningkatkan akurasi hasil secara optimal. *Stopwords* sendiri merujuk pada kumpulan kata yang sering muncul dalam teks namun tidak berdampak signifikan terhadap makna keseluruhan teks. Pada tahap ini, kata-kata yang ada pada dataset akan diidentifikasi dan dihapus dari teks dengan maksud untuk menghilangkan unsur yang tidak relevan atau kurang bermakna dalam analisis sentimen. *Stopword Removal* juga merupakan langkah dalam pemilihan kata dengan mengeluarkan kata-kata yang tidak memberikan kontribusi signifikan dalam data, sehingga dapat meningkatkan akurasi hasil klasifikasi secara optimal.

#### **5. *Stemming***

*Stemming* adalah langkah penghapusan afiks atau imbuhan pada kata dalam suatu kalimat, sehingga kata tersebut disederhanakan menjadi bentuk dasarnya dalam dokumen. Tujuan dari proses *stemming* adalah untuk mengurangi variasi kata yang memiliki akar kata yang sama, dengan demikian meningkatkan kualitas data dan informasi yang ditemukan dalam pencarian informasi. *Stemming* dapat dijelaskan sebagai suatu proses yang mengubah atau mengembalikan suatu kata ke bentuk dasarnya atau memiliki arti yang serupa dengan kata dasar tersebut. *Stemming* melibatkan transformasi kata-kata menjadi bentuk dasarnya, di mana imbuhan di depan atau di belakang kata dihapuskan untuk mencapai hasil tersebut.

### **IV. Augmentasi Data**

Pada tahap ini, dilakukan augmentasi data dengan menambahkan variasi pada data sentimen positif yang ada dalam dataset. Augmentasi data bertujuan untuk mengatasi ketidakseimbangan kelas dalam dataset serta meningkatkan kemampuan model dalam mengenali pola dari berbagai bentuk ekspresi sentimen positif.

Proses augmentasi ini melibatkan teknik-teknik seperti penggantian sinonim atau melakukan parafrasa teks dalam dataset tanpa melakukan perubahan sentimen.

Penambahan variasi ini penting agar model tidak terlalu bergantung pada bentuk kata atau frasa tertentu, melainkan mampu menangkap pola umum dari sentimen positif secara lebih general. Dengan demikian, augmentasi data diharapkan dapat meningkatkan performa model, terutama dalam menghadapi teks-teks yang bervariasi dalam penggunaannya di dunia nyata.

## V. Pembagian Data

Pada tahap ini, data akan dibagi menjadi dua yaitu data latih dan data uji. Data latih digunakan untuk melatih kedua metode, sedangkan data uji digunakan untuk mengevaluasi performa algoritma yang telah dilatih. Pembagian data ini penting untuk menguji kemampuan algoritma dalam mengklasifikasikan sentimen pada data yang belum pernah dilihat sebelumnya.

Pembagian data dilakukan sebanyak 90% sebagai data latih dan 10% sebagai data uji dan pada data latih dilakukan pembagian data dengan ketentuan yang sama berulang sebanyak 10 kali menggunakan *cross validation*. Pembagian data latih dan data uji dapat dilihat pada Tabel 6.

Tabel 6. Pembagian data latih dan data uji

<b>Fold</b>	<b>Data Latih</b>	<b>Data Uji</b>
<b>1</b>	128 - 1280	0-127
<b>2</b>	0 - 127, 256 - 1280	128 -255
<b>3</b>	0 - 255, 384 – 1280	256 – 383
<b>4</b>	0 - 383, 512 - 1280	384 – 511
<b>5</b>	0 - 511, 640 - 1280	512 – 639
<b>6</b>	0 -639, 768 - 1280	640 – 767
<b>7</b>	0 -767, 896 – 1280	768 – 895
<b>8</b>	0 -895, 1024 – 1280	896 – 1023
<b>9</b>	0 -1023, 1152 – 1280	1024 – 1151
<b>10</b>	0 - 1151	1152 – 1280

## VI. Evaluasi

Data yang digunakan untuk analisis sentimen adalah data *tweet* yang telah melalui tahap *preprocessing* dan tahap pembagian data. Setelah dilakukan pembagian data,

tahap selanjutnya adalah melakukan pelatihan kedua metode menggunakan data latih yang telah dibagi sebelumnya. Data tersebut digunakan sebagai input untuk melatih algoritma *Naive Bayes* dan *Support Vector Machine*.

Proses pelatihan kedua metode bertujuan untuk menghasilkan model klasifikasi yang dapat digunakan untuk mengklasifikasikan sentimen pada data yang belum diketahui labelnya.

Pada tahap ini juga menampilkan perbedaan atau persamaan yang akan dilihat dari kedua metode yang dilakukan, yaitu *Naive Bayes* dan *Support Vector Machine* (SVM). Pada tahap pengujian akan melihat perbandingan hasil dari kinerja yang dilakukan dengan algoritma SVM dan *Naive Bayes*. Evaluasi yang digunakan adalah *confusion matrix* yang merepresentasikan prediksi dan kondisi aktual dari data yang dihasilkan oleh algoritma dan *classification report* yang menghitung kualitas prediksi dengan algoritma klasifikasi.

*Confusion Matrix* adalah sebuah metrik yang digunakan untuk menilai kinerja algoritma klasifikasi, yang menunjukkan hasil prediksi dari sistem yang dibandingkan dengan data yang telah dilabeli secara manual. *Confusion Matrix* terdiri dari empat bagian yaitu *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN), yang menggambarkan sejauh mana algoritma mampu mengenali sentimen positif dan negatif dengan benar. Kinerja dari metode klasifikasi umumnya dihitung menggunakan data dalam tabel *confusion matrix* yang dapat dilihat pada Tabel 7.

Tabel 7. *Confusion Matrix*

Fakta	Prediksi	
	Positif	Negatif
Positif	TP ( <i>True Positive</i> )	FN ( <i>False Negative</i> )
Negatif	FP ( <i>False Positive</i> )	TN ( <i>True Negative</i> )

TP merupakan jumlah data yang benar diklasifikasikan sebagai positif, FP merupakan jumlah data yang diklasifikasikan sebagai positif tetapi benar-benar



negatif, TN merupakan jumlah data yang benar diklasifikasikan sebagai negatif, dan FN merupakan jumlah data yang diklasifikasikan sebagai negatif tetapi benar-benar positif (Muludi et al., 2021).

Hasil yang didapat dari *confussion matrix* akan digunakan untuk mengukur kinerja kedua algoritma tersebut dalam melakukan analisis sentimen maskapai penerbangan Indonesia dengan indikator parameter, yaitu :

1. Nilai hasil akurasi (*accuracy*);
2. Nilai presisi (*precicion*);
3. Nilai kemampuan kembali suatu data (*Recall*);
4. *F1-score*

Berikut adalah persamaan dari masing-masing metrix pengukuran dapat dilihat pada persamaan 4,5,6,dan 7.

1. *Accuracy*

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (4)$$

2. *Precision*

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

3. *Recall*

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

4. *F1-Score*

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (7)$$

## VII. Penarikan Kesimpulan

Pada tahap ini, penarikan kesimpulan dilakukan pada penelitian untuk mengetahui suatu kesimpulan yang singkat dan jelas mengenai hasil akhir dari analisis sentimen menggunakan *Naive Bayes dan Support Vector Machine (SVM)*. Kesimpulan ditulis di akhir penelitian dan menjelaskan hasil dari perbandingan antara kedua metode klasifikasi sentimen.

## V. KESIMPULAN DAN SARAN

### 5.1 Kesimpulan

Penelitian yang dilakukan mendapatkan hasil sebagai berikut.

1. Dataset maskapai penerbangan Indonesia berhasil diperoleh melalui media sosial *X*, dengan total 2.098 data yang mencakup empat maskapai penerbangan Indonesia: Garuda (882 data), Lion *Air* (740 data), Batik *Air* (354 data), dan Sriwijaya *Air* (122 data). Data yang terkumpul mencakup sentimen positif dan negatif terhadap masing-masing maskapai. Garuda memiliki 199 data sentimen positif dan 180 sentimen negatif, Lion *Air* mengumpulkan 43 sentimen positif dan 350 sentimen negatif, Batik *Air* memiliki 51 sentimen positif dan 106 sentimen negatif, serta Sriwijaya *Air* dengan 18 sentimen positif dan 45 sentimen negatif.
2. Model *Naive Bayes* dan *Support Vector Machine* telah dibangun untuk melakukan analisis sentimen maskapai penerbangan Indonesia berdasarkan data yang diperoleh melalui media sosial *X*. Kinerja dari kedua model akan dibandingkan melalui hasil evaluasi menggunakan parameter akurasi, presisi, *Recall*, dan *f1-Score* pada *Naive Bayes* dan *Support Vector Machine* dengan pengujian dilakukan pada masing-masing data maskapai serta kombinasi data dari berbagai maskapai. Hasil pengukuran kinerja kedua metode adalah sebagai berikut.

#### A. Garuda

- a. Hasil evaluasi yang didapatkan menggunakan metode *Naive Bayes* dengan pembagian dataset sebesar 90% data latih dan 10% data uji menggunakan *cross validation* mendapatkan rata-rata hasil akurasi sebesar 85%, presisi sebesar 85% untuk data sentimen positif dan 85% untuk data sentimen

negatif, *Recall* sebesar 83% untuk data sentimen positif dan 87% untuk data sentimen negatif, dan *f1-Score* sebesar 84% untuk data sentimen positif dan 86% untuk data sentimen negatif.

- b. Hasil evaluasi yang didapatkan menggunakan metode *Support Vector Machine* dengan pembagian dataset sebesar 90% data latih dan 10% data uji menggunakan *cross validation* mendapatkan hasil akurasi sebesar 85%, presisi sebesar 86% untuk data sentimen positif dan 84% untuk data sentimen negatif, *Recall* sebesar 82% untuk data sentimen positif dan 87% untuk data sentimen negatif, dan *f1-Score* sebesar 84% untuk data sentimen positif dan 86% untuk data sentimen negatif.

Berdasarkan hasil tersebut dapat dikatakan bahwa metode *Support Vector Machine* memiliki kinerja yang seimbang dengan metode *Naive Bayes* dalam melakukan klasifikasi sentimen maskapai penerbangan Garuda.

#### B. Lion Air

- a. Hasil evaluasi yang didapatkan menggunakan metode *Naive Bayes* dengan pembagian dataset sebesar 90% data latih dan 10% data uji menggunakan *cross validation* mendapatkan rata-rata hasil akurasi sebesar 97%, presisi sebesar 95% untuk data sentimen positif dan 99% untuk data sentimen negatif, *Recall* sebesar 99% untuk data sentimen positif dan 94% untuk data sentimen negatif, dan *f1-Score* sebesar 97% untuk data sentimen positif dan 96% untuk data sentimen negatif.
- b. Hasil evaluasi yang didapatkan menggunakan metode *Support Vector Machine* dengan pembagian dataset sebesar 90% data latih dan 10% data uji menggunakan *cross validation* mendapatkan hasil akurasi sebesar 98%, presisi sebesar 98% untuk data sentimen positif dan 98% untuk data sentimen negatif, *Recall* sebesar 98% untuk data sentimen positif dan 98% untuk data sentimen negatif, dan *f1-Score* sebesar 98% untuk data sentimen positif dan 98% untuk data sentimen negatif.

Berdasarkan hasil tersebut dapat dikatakan bahwa metode *Support Vector Machine* memiliki kinerja yang lebih baik dibandingkan metode *Naive Bayes* dalam melakukan klasifikasi analisis sentimen maskapai penerbangan Lion Air.

### C. Batik Air

- a. Hasil evaluasi yang didapatkan menggunakan metode *Naïve Bayes* dengan pembagian dataset sebesar 90% data latih dan 10% data uji menggunakan *cross validation* mendapatkan rata-rata hasil akurasi sebesar 85%, presisi sebesar 97% untuk data sentimen positif dan 79% untuk data sentimen negatif, *Recall* sebesar 73% untuk data sentimen positif dan 97% untuk data sentimen negatif, dan *f1-Score* sebesar 82% untuk data sentimen positif dan 87% untuk data sentimen negatif.
- b. Hasil evaluasi yang didapatkan menggunakan metode *Support Vector Machine* dengan pembagian dataset sebesar 90% data latih dan 10% data uji menggunakan *cross validation* mendapatkan hasil akurasi sebesar 90%, presisi sebesar 90% untuk data sentimen positif dan 92% untuk data sentimen negatif, *Recall* sebesar 93% untuk data sentimen positif dan 87% untuk data sentimen negatif, dan *f1-Score* sebesar 91% untuk data sentimen positif dan 89% untuk data sentimen negatif.

Berdasarkan hasil tersebut dapat dikatakan bahwa metode *Support Vector Machine* memiliki kinerja yang lebih baik dibandingkan metode *Naïve Bayes* dalam melakukan klasifikasi analisis sentimen maskapai penerbangan Batik Air.

### D. Sriwijaya Air

- a. Hasil evaluasi yang didapatkan menggunakan metode *Naïve Bayes* dengan pembagian dataset sebesar 90% data latih dan 10% data uji menggunakan *cross validation* mendapatkan rata-rata hasil akurasi sebesar 91%, presisi sebesar 98% untuk data sentimen positif dan 87% untuk data sentimen negatif, *Recall* sebesar 82% untuk data sentimen positif dan 99% untuk data sentimen negatif, dan *f1-Score* sebesar 89% untuk data sentimen positif dan 92% untuk data sentimen negatif.
- b. Hasil evaluasi yang didapatkan menggunakan metode *Support Vector Machine* dengan pembagian dataset sebesar 90% data latih dan 10% data uji menggunakan *cross validation* mendapatkan hasil akurasi sebesar 94%, presisi sebesar 93% untuk data sentimen positif dan 95% untuk data sentimen negatif, *Recall* sebesar 95% untuk data sentimen positif dan 94%

untuk data sentimen negatif, dan *f1-Score* sebesar 94% untuk data sentimen positif dan 95% untuk data sentimen negatif.

Berdasarkan hasil tersebut dapat dikatakan bahwa metode *Support Vector Machine* memiliki kinerja yang lebih baik dibandingkan metode *Naive Bayes* dalam melakukan klasifikasi analisis sentimen maskapai penerbangan Garuda.

#### E. Kombinasi Data Maskapai Penerbangan Indonesia

- a. Hasil evaluasi yang didapatkan menggunakan metode *Naive Bayes* dengan pembagian dataset sebesar 90% data latih dan 10% data uji menggunakan *cross validation* mendapatkan rata-rata hasil akurasi sebesar 89%, presisi sebesar 90% untuk data sentimen positif dan 88% untuk data sentimen negatif, *Recall* sebesar 87% untuk data sentimen positif dan 92% untuk data sentimen negatif, dan *f1-Score* sebesar 89% untuk data sentimen positif dan 89% untuk data sentimen negatif.
- b. Hasil evaluasi yang didapatkan menggunakan metode *Support Vector Machine* dengan pembagian dataset sebesar 90% data latih dan 10% data uji menggunakan *cross validation* mendapatkan hasil akurasi sebesar 92%, presisi sebesar 96% untuk data sentimen positif dan 89% untuk data sentimen negatif, *Recall* sebesar 88% untuk data sentimen positif dan 96% untuk data sentimen negatif, dan *f1-Score* sebesar 92% untuk data sentimen positif dan 92% untuk data sentimen negatif.

Berdasarkan hasil tersebut dapat dikatakan bahwa metode *Support Vector Machine* memiliki kinerja yang lebih baik dibandingkan metode *Naive Bayes* dalam melakukan klasifikasi analisis sentimen maskapai penerbangan Indonesia.

## 5.2 Saran

Beberapa saran yang dapat diberikan untuk penelitian selanjutnya adalah sebagai berikut.

1. Penggunaan metode klasifikasi lainnya seperti *Deep Learning* untuk mendapatkan nilai akurasi yang lebih baik dari penelitian ini dan melakukan

analisis lebih dalam seperti mempertimbangkan kelas netral, dan mengelompokkan tingkatan sentimen pada masing-masing kelas.

2. Penambahan data pada dataset untuk mendapatkan akurasi yang lebih baik seperti mengambil data dari *Youtube*, *Instagram*, *Facebook*, dan lain-lain. Model *machine learning* membutuhkan data latih yang banyak untuk dataset yang cukup kompleks.
3. Pembangunan sistem yang dapat melakukan klasifikasi sentimen berdasarkan model yang telah dibangun.

## DAFTAR PUSTAKA

- Airin, K., Farhanah, N., Hajar, H., & Septem, L. (2022). The Best Malaysian Airline Companies Visualization through Bilingual Twitter Sentiment Analysis: A Machine Learning Classification. In *International Journal on Informatics Visualization* (Vol. 6, Issue 1, pp. 130–137).
- Azmarani, A. (2016). Kebijakan dan Manajemen Publik Analisis Kualitas Pelayanan Maskapai Penerbangan Low Cost Carrier (Studi Deskriptif Di PT. Citilink Indonesia Cabang Surabaya). In *Jurnal Kebijakan dan Manajemen Publik* (Vol. 3, Issue 4, pp. 1–6).
- Evgeniou, T., & Pontil, M. (2001). *Support Vector Machines: Theory and applications. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2049 LNAI, 249–257. [https://doi.org/10.1007/3-540-44673-7\\_12](https://doi.org/10.1007/3-540-44673-7_12)
- Fikri, M., Sabrila, T., & Azhar, Y. (2020). Perbandingan Metode Naïve Bayes dan *Support Vector Machine* pada Analisis Sentimen Twitter. In *SMATIKA JURNAL* (Vol. 10, Issue 02, pp. 71–76). Sekolah Tinggi Informatika dan Komputer Indonesia. <https://doi.org/10.32664/smatika.v10i02.455>
- Husada, H., & Paramita, A. (2021). Analisis Sentimen Pada Maskapai Penerbangan di Platform Twitter Menggunakan Algoritma *Support Vector Machine* (SVM). In *Teknika* (Vol. 10, Issue 1, pp. 18–26). Institut Informatika Indonesia Surabaya.
- Jadhav, S., & Channe, H. (2016). Comparative Study of K-NN, *Naive Bayes* and Decision Tree Classification Techniques. In *International Journal of Science and Research (IJSR)* (Vol. 5, Issue 1, pp. 1842–1845).

- Muludi, K., Akbar, M., Shofiana, D., & Syarif, A. (2021). Sentiment Analysis Of Energy Independence Tweets Using Simple Recurrent Neural Network. In *Indonesian Journal of Computing and Cybernetics Systems (IJCCS)* (Vol. 15, Issue 4, p. 339). Universitas Gadjah Mada.
- Nahumury, A., Manongga, D., & Iriani, A. (2022). Analysis Sentiment On Airline Customer Saisfaction Using Reccurent Neural Network. In *Eduvest-Journal of Universal Studies* (Vol. 2, Issue 10, pp. 2120–2128).
- Nurdina, A., & Puspita, A. (2023). *Naive Bayes* and KNN for Airline Passenger Satisfaction Classification: Comparative Analysis. In *Journal of Information System Exploration and Research* (Vol. 1, Issue 2, pp. 83–92). Surya Hijau Manfaat. <https://doi.org/10.52465/joiser.v1i2.167>
- Pravina, A., Cholissodin, I., & Adikara, P. (2019). Analisis Sentimen Tentang Opini Maskapai Penerbangan pada Dokumen Twitter Menggunakan Algoritme *Support Vector Machine* (SVM). In *Jurnal Pengembangan Teknologi dan Ilmu Komputer* (Vol. 3, Issue 3, pp. 2789–2797).
- Prissia, D., & Daryanto, W. (2019). Financial Performance Analysis and Evaluation of Airline Industry Indonesia: Case Study of PT Garuda Indonesia to Support Vision 2020 “BEYOND THE SKY” for The Period of 2014-2018. In *International Journal of Business, Economics and Law* (Vol. 20, Issue 1, pp. 10–21).
- Ruz, G. A., Henríquez, P. A., & Mascareño, A. (2020). Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers. In *Future Generation Computer Systems* (Vol. 106, Issue 6, pp. 92–104). <https://doi.org/10.1016/j.future.2020.01.005>
- Sahayak, V., Shete, V., & Pathan, A. (2015). Sentiment Analysis on Twitter Data. In *International Journal of Innovative Research in Advanced Engineering (IJIRAE)* (Vol. 1, Issue 2, pp. 178–183).
- Santoso, H., Armansyah, A., & Desliani, D. (2022). Analisis Sentimen Mahasiswa Terkait Pembelajaran Tatap Muka Menggunakan Metode *Naive Bayes* Classifier. *Techno.Com* (Vol. 21, Issue 3, pp. 644–654).



Sari, F., & Wibowo, A. (2019). Analisis Sentimen Pelanggan Toko Online JD.ID Menggunakan Metode *Naive Bayes* Classifier Berbasis Konversi Ikon Emosi. In *Jurnal SIMETRIS* (Vol. 10, Issue 2, pp. 681–686).

Srinath, R. (2017). The Fastest Growing Programming Language. In *International Research Journal of Engineering and Technology (IRJET)* (Vol. 4, Issue 12, pp. 354–357). STMIK Sinar Nusantara Surakarta.

Uddin, M., Hafi, F., Hossain, S., & Mohammad, S. (2022). Drug Sentiment Analysis using Machine Learning Classifiers. In *International Journal of Advanced Computer Science and Applications* (Vol. 13, Issue 1, pp. 92–100). Science and Information Organization.