

**Perbandingan Metode Support Vector Machine Dan Naïve Bayes
Pada Analisis Sentimen Maskapai Lion Air Pada Media Sosial Twitter**



SKRIPSI

**Disusun untuk memenuhi sebagian persyaratan
mencapai derajat Sarjana**



Disusun Oleh:

Sindi Rella
2000018156

**PROGRAM STUDI S1 INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS AHMAD DAHLAN
2024**

**PERBANDINGAN METODE SUPPORT VECTOR MACHINE DAN NAÏVE BAYES
PADA ANALISIS SENTIMEN MASKAPAI LION AIR PADA MEDIA SOSIAL
TWITTER
SKRIPSI**



Disusun Oleh:

SINDI RELLA
2000018156

**PROGRAM STUDI S1 INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS AHMAD DAHLAN**

2024

LEMBAR PERSETUJUAN PEMBIMBING

SKRIPSI

**PERBANDINGAN METODE SUPPORT VECTOR MACHINE DAN NAÏVE
BAYES PADA ANALISIS SENTIMEN MASKAPAI LION AIR PADA MEDIA
SOSIAL TWITTER**

Dipersiapkan dan disusun oleh:

**SINDI RELLA
2000018156**

**Program Studi S1 Informatika
Fakultas Teknologi Industri
Universitas Ahmad Dahlan**

Telah disetujui oleh:

Pembimbing



Miftahurrahma Rosyda, S.Kom., M.Eng.

19900615 201908 011 1029280

LEMBAR PENGESAHAN

SKRIPSI

PERBANDINGAN METODE SUPPORT VECTOR MACHINE DAN NAÏVE BAYES PADA ANALISIS SENTIMEN MASKAPAI LION AIR PADA MEDIA SOSIAL TWITTER


Dipersiapkan dan disusun oleh:

SINDI RELLA
2000018156


Telah dipertahankan di depan Dewan Penguji
pada Selasa, 13 Agustus 2024
dan dinyatakan telah memenuhi syarat

Susunan Dewan Penguji


Ketua : Miftahurrahma Rosyda, S.Kom., M.Eng.

 13/08/24

Penguji 1 : Herman Yuliansyah, S.T., M.Eng., Ph.D.

 13/08 24.

Penguji 2 : Dewi Soyusiawaty, S.T., M.T.

 . 27/8/2024

Yogyakarta, Senin 8 Juli 2024

Dekan Fakultas Teknologi Industri
Universitas Ahmad Dahlan




Prof. Dr. Ir. Siti Jamilatun, M.T.
19660812 199601 011 0784324

LEMBAR PERNYATAAN KEASLIAN

SURAT PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Sindi Rella
NIM : 2000018156
Prodi : Informatika
Judul TA/Skripsi : PERBANDINGAN PERFORMA METODE SUPPORT VECTOR MACHINE
DAN NAÏVE BAYES PADA ANALISIS SENTIMEN MASKAPAI LION AIR

Dengan ini saya menyatakan bahwa Laporan Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar Ahli Madya/Kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 08 Juli 2024

Mengetahui,
Dosen Pembimbing



Miftahurrahma Rosyda, S.Kom., M.Eng.
19900615 201908 011 1029280

Yang menyatakan,



Sindi Rella
2000018156

PERNYATAAN PERSETUJUAN AKSES

Saya yang bertanda tangan di bawah ini:

Nama : Sindi Rella

NIM : 2000018156 Email : sindirella2000018156@webmail.uad.ac.id

Fakultas : Teknologi Industri Program Studi : Informatika

Judul tugas akhir : Perbandingan Metode Support Vector Machine Dan Naïve Bayes Pada Analisis Sentimen Maskapai Lion Air Pada Media Sosial Twitter

Dengan ini saya menyerahkan hak *sepenuhnya* kepada Perpustakaan Universitas Ahmad Dahlan untuk menyimpan, mengatur akses serta melakukan pengelolaan terhadap karya saya ini dengan mengacu pada ketentuan akses tugas akhir elektronik sebagai berikut

Saya (~~mengijinkan~~/~~tidak mengijinkan~~)* karya tersebut diunggah ke dalam Repository Perpustakaan Universitas Ahmad Dahlan.

Demikian pernyataan ini saya buat dengan sebenarnya.

Yogyakarta, 27 Agustus 2024



Sindi Rella

Mengetahui,
Pembimbing**



Miftahurrahma Rosyda, S.Kom., M.Eng.
NIP/NIPM.19900615 201908 011 1029280

Ket:

*coret salah satu

**jika diijinkan TA dipublish maka ditandatangani dosen pembimbing dan mahasiswa

KATA PENGANTAR

Puji dan syukur kita panjatkan kehadirat Allah SWT tuhan Maha Pencipta dan Maha Kuasa yang telah memberikan kita hidayah serta karunia-Nya sehingga dapat menyelesaikan skripsi ini sebagai salah satu persyaratan dalam mencapai gelas Sarjana Strata Satu di Fakultas Teknologi Industri Program Studi Informatika Universitas Ahmad Dahlan dengan judul **“PERBANDINGAN PERFORMA METODE SUPPORT VECTOR MACHINE DAN NAÏVE BAYES PADA ANALISIS SENTIMEN MASKAPAI LION AIR PADA MEDIA SOSIAL TWITTER”**. Tak lupa shalawat dan salam kita curahkan kepada junjungan kita Nabi Muhammad SAW yang telah menuntun kita dari zaman jahiliyah menuju zaman keislaman ini.

Dalam penyusunan Skripsi ini, banyak dukungan serta bantuan tenaga, materi, informasi, dan waktu yang tak terhingga telah diterima dari berbagai pihak. Menyadari bahwa Skripsi ini tidak akan terwujud dengan baik tanpa dukungan dari orang-orang sekitar, maka dari itu rasa terima kasih yang mendalam ingin disampaikan kepada:

1. Bapak Prof. Dr. Muchlas Arkanuddin, M.T. selaku Rektor Universitas Ahmad Dahlan.
2. Ibu Prof. Dr. Ir. Siti Jamilatun, M.T. selaku Dekan Fakultas Teknologi Industri Universitas Ahmad Dahlan.
3. Bapak Dr. Murinto, S.Si.,M.Kom.,selaku Kepala Program Studi Informatika, Universitas Ahmad Dahlan.
4. Ibu Miftahurrahma Rosyda, S.Kom., M.Eng. selaku Dosen Pembimbing Skripsi.
5. Ibu Murein Miksa Mardhia S.T, M.T sebagai Dosen Pembimbing Akademik.
6. Bapak dan ibu Dosen Informatika Universitas Ahmad Dahlan Yogyakarta yang telah senantiasa memberikan ilmunya sehingga Skripsi ini dapat selesai dengan baik.
7. Kedua orang tua yang senantiasa memberikan doa serta dukungannya selama ini.
8. Teman seperjuangan saya Fauziah dan Silfa yang berjuang bersama selama 4 tahun di bangku perkuliahan, serta teman-teman saya Patar, Rosie, Pirja, Brayen, Ikbal, Egi yang selalu bersedia menjadi tempat berkeluh kesah.

Sadar bahwa Skripsi ini masih jauh dari kata sempurna, diharapkan kritik dan saran yang membangun dari para pembaca. Harapannya, semoga Skripsi ini dapat bermanfaat, baik bagi diri sendiri maupun bagi para pembaca yang akan melakukan penelitian lebih lanjut.

DAFTAR ISI

| | |
|--|------|
| HALAMAN JUDUL | i |
| KATA PENGANTAR..... | vi |
| DAFTAR ISI..... | vii |
| DAFTAR GAMBAR..... | viii |
| DAFTAR TABEL..... | ix |
| DAFTAR KODE PROGRAM..... | x |
| DAFTAR SINGKATAN DAN ARTI LAMBANG | xi |
| ABSTRAK..... | xii |
| BAB I Pendahuluan | 1 |
| 1.1. Latar Belakang Masalah | 1 |
| 1.2. Batasan Masalah Penelitian | 3 |
| 1.3. Rumusan Masalah | 4 |
| 1.4. Tujuan Penelitian..... | 4 |
| 1.5. Manfaat Penelitian | 4 |
| BAB II Tinjauan Pustaka | 5 |
| 2.1. Kajian Penelitian Terdahulu | 5 |
| 2.2. Landasan Teori | 8 |
| BAB III Metodologi Penelitian..... | 27 |
| 3.1. Metode Pengumpulan Data | 27 |
| 3.2. Alat Penelitian | 27 |
| 3.3. Tahapan Penelitian..... | 28 |
| 3.4. Pengujian Akurasi..... | 31 |
| BAB IV Hasil dan Pembahasan | 32 |
| 4.1. Hasil Pengumpulan Data | 32 |
| 4.2. Tahapan Penelitian..... | 33 |
| 4.3. Pengujian Tanpa SMOTE | 45 |
| 4.4. Pengujian Akurasi Dengan SMOTE..... | 46 |
| BAB V Kesimpulan dan Saran | 51 |
| 5.1. Kesimpulan..... | 51 |
| 5.2. Saran..... | 52 |
| DAFTAR PUSTAKA..... | 53 |
| LAMPIRAN | 57 |

DAFTAR GAMBAR

| | |
|---|----|
| Gambar 2. 1 <i>Confusion Matrix</i> | 14 |
| Gambar 3. 1 <i>Flowchart</i> | 28 |
| Gambar 3. 2 <i>Flowchart</i> Pengolahan Data | 29 |
| Gambar 4. 1 Visualisasi | 43 |
| Gambar 4. 2 <i>Most Positive Tweets</i> | 44 |
| Gambar 4. 3 <i>Most Negative Tweets</i> | 44 |
| Gambar 4. 4 <i>Most Neutral Tweets</i> | 45 |
| Gambar 4. 5 Hasil <i>Oversampling</i> | 47 |
| Gambar 4. 6 Penerapan <i>SMOTE Oversampling</i> | 47 |
| Gambar 4. 7 <i>Confusion Matrix SVM</i> | 49 |
| Gambar 4. 8 <i>Confusion Matrix NB</i> | 50 |

DAFTAR TABEL

| | |
|--|----|
| Tabel 2. 1 Perbandingan beberapa penelitian terdahulu mengenai klasifikasi | 7 |
| Tabel 2. 2 Contoh Studi Kasus | 15 |
| Tabel 2. 3 Hasil <i>Term</i> Dokumen | 18 |
| Tabel 4. 1 Hasil Pengumpulan Data..... | 32 |
| Tabel 4. 2 Hasil <i>Cleaning</i> | 34 |
| Tabel 4. 3 Hasil <i>Translate</i> | 35 |
| Tabel 4. 4 Hasil <i>Case Folding</i> | 36 |
| Tabel 4. 5 Hasil Tokenisasi..... | 37 |
| Tabel 4. 6 Hasil <i>Stopword Removal</i> | 38 |
| Tabel 4. 7 Hasil <i>Stemming</i> | 39 |
| Tabel 4. 8 Hasil Label VADER 1 | 39 |
| Tabel 4. 9 Pengujian VADER 1 | 40 |
| Tabel 4. 10 Hasil Label VADER 2 | 40 |
| Tabel 4. 11 Pengujian VADER 2 | 40 |
| Tabel 4. 12 Hasil Label VADER 3 | 41 |
| Tabel 4. 13 Pengujian VADER 3 | 41 |
| Tabel 4. 14 Hasil <i>Labeling</i> | 42 |
| Tabel 4. 15 Perbandingan Klasifikasi SVM dan NB..... | 43 |
| Tabel 4. 16 Perbandingan Hasil Tanpa SMOTE. | 46 |
| Tabel 4. 17 <i>Confusion Matrix</i> SVM | 48 |
| Tabel 4. 18 <i>Confusion Matrix</i> NB..... | 50 |
| Tabel 4. 19 Hasil Perbandingan | 50 |

DAFTAR KODE PROGRAM

| | |
|--|----|
| Kode Program 4. 1 <i>Crawling Data</i> | 33 |
| Kode Program 4. 2 <i>Cleaning Data</i> | 34 |
| Kode Program 4. 3 <i>Translate</i> | 35 |
| Kode Program 4. 4 <i>Case Folding</i> | 36 |
| Kode Program 4. 5 <i>Tokenisasi</i> | 36 |
| Kode Program 4. 6 <i>Stopword Removal</i> | 38 |
| Kode Program 4. 7 <i>Stemming</i> | 39 |
| Kode Program 4. 8 <i>Polarity Score</i> | 41 |
| Kode Program 4. 9 <i>Labeling</i> | 42 |
| Kode Program 4. 10 <i>TFIDF</i> | 45 |
| Kode Program 4. 11 <i>Split Dataset</i> | 46 |
| Kode Program 4. 12 <i>SVM Tanpa SMOTE</i> | 46 |
| Kode Program 4. 13 <i>NB Tanpa SMOTE</i> | 46 |
| Kode Program 4. 14 <i>SMOTE Oversampling</i> | 47 |
| Kode Program 4. 15 <i>Support Vector Machine</i> | 48 |
| Kode Program 4. 16 <i>Naive Bayes</i> | 49 |

DAFTAR SINGKATAN DAN ARTI LAMBANG

| | |
|--------|---|
| SVM | : <i>Support Vector Machine</i> |
| NB | : <i>Niave Bayes</i> |
| SMOTE | : <i>Synthetic Minority Over-sampling Technique</i> |
| TF-IDF | : <i>Pembobotan Term Frequency-Inverse Document Frequency</i> |

ABSTRAK

Saat ini perkembangan teknologi informasi telah menjadikan sosial media menjadi salah satu sarana yang paling banyak digunakan masyarakat untuk mendapatkan informasi serta menuangkan opini dan pendapat sebagai salah satu bentuk dari respon atau tanggapan sekelompok masyarakat mengenai suatu permasalahan. Analisis sentimen telah memberikan wawasan mengenai perspsi atau opini masyarakat diberbagai macam aspek seperti transportasi. Pesawat telah menjadi alat transportasi antar pulau dan antar kota yang paling banyak digunakan masyarakat indonesia sehingga penelitian ini berfokus pada analisis sentimen terhadap cuitan masyarakat mengenai maskapai lion air di media sosial Twitter dengan tujuan untuk melihat komentar publik mengenai maskapai penerbangan Lion Air.

Pada penelitian ini digunakan dataset sebanyak 2000 cuitan dengan tahapan *preprocessing* seperti *case folding*, *stemming* dan *stopword removal* yang membuat dataset sesuai dengan standarisasi. Dilakukan analisis sentimen terhadap cuitan yang berkaitan dengan maskapai Lion Air dengan tujuan untuk mengetahui opini dari masyarakat mengenai maskapai penerbangan Lion Air dengan menggunakan metode klasifikasi *Support Vector Machine* dan *Naïve Bayes*. Pengujian ini dilakukan untuk menemukan hasil dari setiap metode yang kemudian dilakukan perbandingan dari kedua hasil metode *Support Vector Machine* dan *Naïve Bayes*

Dataset dibagi menjadi dua sesuai dengan kebutuhan penelitian dimana terdapat data uji dan data latih masing masing sebanyak 20% dan 80%, sehingga penelitian ini akan menghasilkan perbandingan hasil akurasi dengan metode *Support Vector Machine* sebesar 80% dan metode *Naïve Bayes* sebesar 74% dengan menggunakan teknik SMOTE. Hasil tersebut membuktikan bahwa metode *Support Vector Machine* lebih unggul dalam menangani persoalan mengenai maskapai Lion Air.

Keywords: Lion Air, *Naïve Bayes*, *Sentiment Analysis*, *Support Vector Machine*, Twitter