

PERBANDINGAN METODE *NAÏVE BAYES* DAN
SUPPORT VECTOR MACHINE (SVM) PADA
ANALISIS SENTIMEN

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh:

FENNY UTARY
NIM: 09021181924001

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2023

LEMBAR PENGESAHAN SKRIPSI

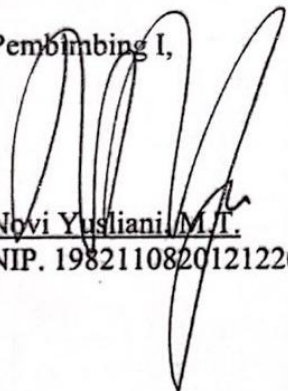
PERBANDINGAN METODE NAÏVE BAYES DAN SUPPORT VECTOR MACHINE (SVM) PADA ANALISIS SENTIMEN

Oleh :

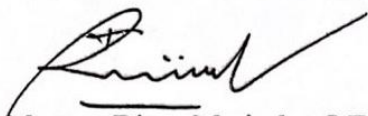
FENNY UTARY
NIM: 09021181924001

Palembang, 28 Juli 2023



Pembimbing I,


Novi Yusliani, M.T.
NIP. 198211082012122001

Pembimbing II,


Mastura Diana Marieska, S.T., M.T.
NIP. 198603212018032001

Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003

TANDA LULUS UJIAN SIDANG SKRIPSI

Pada hari Kamis tanggal 27 Juli 2023 telah dilaksanakan ujian komprehensif skripsi oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Fenny Utary
NIM : 09021181924001
Judul : Perbandingan Metode Naïve Bayes dan Support Vector Machine (SVM)
Pada Analisis Sentimen

Dan dinyatakan LULUS

1. Ketua Penguji

Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003



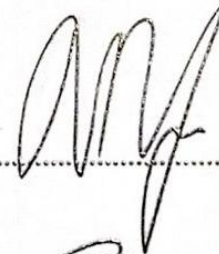
2. Penguji

Kanda Januar Miraswan, M.T.
NIP. 199001092019031012



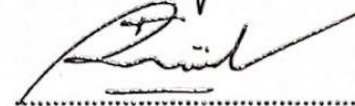
3. Pembimbing I

Novi Yusliani, M.T.
NIP. 198211082012122001



4. Pembimbing II

Mastura Diana Marieska, S.T., M.T.
NIP. 198603212018032001



Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Fenny Utary
NIM : 09021181924001
Program Studi : Teknik Informatika
Judul Skripsi : Perbandingan Metode Naïve Bayes dan Support Vector
Machine (SVM) Pada Analisis Sentimen

Hasil Pengecekan Software *iThenticate/Turnitin* : 6%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan tugas akhir ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya.



Palembang, 28 Juli 2023



Fenny Utary
NIM. 09021181924001

Motto dan Persembahan

“Tidak ada sesuatu yang mustahil untuk dikerjakan,

Hanya tidak ada sesuatu yang mudah.”

- Napoleon Bonaparte -

Ku persembahkan karya tulis ini kepada:

- Kedua Orang Tua
- Adik Kandungku
- Keluarga Besar
- Para Dosen
- Teman Hidup
- Teman di Teknik Informatika
- Fakultas Ilmu Komputer
- Universitas Sriwijaya

ABSTRACT

The study aims to perform a sentiment analysis in user review data to find out user satisfaction in the Spotify apps. Sentiment analysis is the process of document classification and in this study is divided into two parts, namely positive class and negative class. Classification is very important in searching documents for developers. The classification process begins by dividing the data collection into training data and testing data. Training data used Naïve Bayes method and Support Vector Machine method to obtain classification model for class determination on testing data. Both of the method is used to classify the results of user review data written on review section of Spotify apps so as to produced the desired classification automatically. The results of this study shows that Support Vector Machine gets better score with an accuracy of 83%, precision of 89%, recall of 64%, and f-measure of 74% than Naïve Bayes with an accuracy of 61%, precision of 57%, recall of 43%, and f-measure of 49%.

Keywords : Sentiment Analysis, Comparative Study, Naïve Bayes, Support Vector Machine, Spotify Apps

ABSTRAK

Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap data ulasan pengguna untuk mengetahui kepuasan pengguna terhadap aplikasi *Spotify*. Analisis sentimen adalah proses klasifikasi dokumen dan dalam penelitian ini dibagi menjadi dua bagian yaitu kelas positif dan kelas negatif. Klasifikasi sangat penting dalam pencarian dokumen bagi pengembang. Proses klasifikasi diawali dengan membagi koleksi data menjadi data latih dan data uji. Data latih digunakan metode *Naïve Bayes* dan metode *Support Vector Machine* sehingga diperoleh model klasifikasi untuk menentukan kelas pada data uji. Dua metode ini digunakan untuk mengklasifikasikan hasil data ulasan pengguna yang dituliskan pada kolom ulasan terhadap aplikasi *Spotify* sehingga menghasilkan klasifikasi secara otomatis yang diinginkan. Hasil dari penelitian ini menunjukkan bahwa metode *Support Vector Machine* memiliki nilai lebih unggul dengan nilai akurasi sebesar 83%, presisi sebesar 89%, *recall* sebesar 64%, dan *f-measure* sebesar 74%, pada metode *Naïve Bayes* dihasilkan nilai akurasi sebesar 61%, presisi sebesar 57%, *recall* sebesar 43%, dan *f-measure* sebesar 49%.

Kata Kunci : Analisis Sentimen, Studi Perbandingan, *Naïve Bayes*, *Support Vector Machine*, Aplikasi *Spotify*

KATA PENGANTAR

Puji dan syukur atas berkat dan nikmat yang telah Allah swt. berikan kepada penulis sehingga mampu menyelesaikan Skripsi dengan judul “Perbandingan Metode Naïve Bayes dan Support Vector Machine Pada Analisis Sentimen”. Tujuan dari penulisan skripsi ini adalah untuk memperoleh gelar Sarjana Komputer di Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Penulis sadar bahwa penulis tidak akan mampu menyelesaikan proses penulisan skripsi ini tanpa bantuan orang lain. Dalam menyelesaikan kegiatan ini banyak pihak yang memberikan bantuan dan dukungan. Oleh karena itu, penulis menghaturkan banyak terima kasih kepada:

1. Allah swt. yang telah memberikan keberkahan dan meridhai setiap langkah sehingga penulis dapat menyelesaikan dengan baik.
2. Orang tua dan keluarga yang telah memberikan dukungan moral maupun materil.
3. Ibu Alvi Syahrini Utami, M.Kom. selaku Ketua Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
4. Bapak Dr. Ali Ibrahim, M.T. selaku Dosen Pembimbing Akademik penulis.
5. Ibu Novi Yusliani, M.T. dan Ibu Mastura Diana Marieska, S.T., M.T. selaku Dosen Pembimbing Skripsi penulis.

6. Seluruh dosen, Pak Ricy, dan staff Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.
7. Azza, Julia, Faliha, Vinito, Theja, Farhan, dan Aditya yang telah memberikan semangat.
8. Zafira, Tarisa, Nilam, Heni, Fidyah, Nadia, Nabila, Erish, Nancy dan seluruh teman-teman Teknik Informatika 2019.

Penulis sadar akan adanya kekurangan dalam penulisan Skripsi ini. Oleh karena itu, penulis menerima jika terdapat saran dan kritik yang konstruktif. Semoga skripsi ini dapat memberi kebermanfaatan bagi banyak orang.

Palembang, 28 Juli 2023

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERSETUJUAN KOMISI PENGUJI.....	iii
HALAMAN PERNYATAAN.....	iv
ABSTRACT.....	v
ABSTRAK.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xv
BAB I PENDAHULUAN	I-1
1.1 Pendahuluan	I-1
1.2 Latar Belakang	I-1
1.3 Rumusan Masalah	I-4
1.4 Tujuan Penelitian	I-4
1.5 Manfaat Penelitian	I-5
1.6 Batasan Masalah	I-5
1.7 Sistematika Penulisan	I-5
1.8 Kesimpulan	I-7
BAB II KAJIAN LITERATUR	II-1
2.1 Pendahuluan	II-1
2.2 Analisis Sentimen	II-1

2.3 <i>Text Preprocessing</i>	I-2
2.4 TF-IDF (<i>Term Frequency Inverse Document Frequency</i>)	II-4
2.5 <i>Negation Handling</i>	II-6
2.6 <i>Naïve Bayes</i>	II-6
2.7 <i>Support Vector Machine (SVM)</i>	II-7
2.8 <i>Confusion Matrix</i>	II-8
2.9 <i>Rational Unified Process (RUP)</i>	II-10
2.10 Penelitian Lain yang Relevan	II-11
2.11 Kesimpulan	II-14
BAB III METODOLOGI PENELITIAN	III-1
3.1 Pendahuluan	III-1
3.2 Pengumpulan Data	III-1
3.3 Tahapan Penelitian	III-1
3.3.1 Kerangka Kerja	III-2
3.3.2 Kriteria Pengujian	III-5
3.3.3 Format Data Pengujian	III-5
3.3.4 Alat yang Digunakan dalam Pelaksanaan Penelitian	III-6
3.3.5 Pengujian Penelitian	III-7
3.3.6 Analisis Hasil Pengujian dan Membuat Kesimpulan	III-7
3.4 Metode Pengembangan Perangkat Lunak	III-7
3.4.1 Fase Insepsi	III-8
3.4.2 Fase Elaborasi	III-8
3.4.3 Fase Konstruksi	III-8
3.4.4 Fase Transisi	III-9
3.5 Manajemen Proyek Penelitian	III-9
3.6 Kesimpulan	III-14

BAB IV PENGEMBANGAN PERANGKAT LUNAK.....	IV-1
4.1 Pendahuluan	IV-1
4.2 Fase Insepsi	IV-1
4.2.1 Pemodelan Bisnis	IV-1
4.2.2 Kebutuhan Sistem.....	IV-2
4.2.3 Analisis dan Desain	IV-3
4.3 Fase Elaborasi	IV-30
4.3.1 Pemodelan Bisnis	IV-30
4.3.2 Perancangan Data	IV-30
4.3.3 Perancangan Antarmuka.....	IV-31
4.3.4 Kebutuhan Sistem.....	IV-33
4.3.5 Diagram Aktivitas	IV-33
4.3.6 Diagram <i>Sequence</i>	IV-36
4.4 Fase Konstruksi.....	IV-37
4.4.1 Kebutuhan Sistem.....	IV-38
4.4.2 Diagram Kelas	IV-38
4.4.3 Implementasi	IV-38
4.5 Fase Transisi	IV-40
4.5.1 Pemodelan Bisnis	IV-41
4.5.2 Rencana Pengujian	IV-41
4.5.3 Implementasi	IV-43
4.6 Kesimpulan	IV-46
 BAB V HASIL DAN ANALISIS PENELITIAN.....	 V-1
5.1 Pendahuluan	V-1
5.2 Data Hasil Penelitian.....	V-1
5.2.1 Konfigurasi Percobaan	V-1
5.2.2 Hasil Pengujian Metode NB Pada <i>Data Review Spotify</i>	V-2

5.2.3 Hasil Pengujian Metode SVM Pada <i>Data Review Spotify</i>	V-3
5.3 Analisis Hasil Pengujian Secara Keseluruhan	V-6
5.4 Kesimpulan	V-7
BAB VI KESIMPULAN DAN SARAN	VI-1
6.1 Kesimpulan	VI-1
6.2 Saran	VI-1
DAFTAR PUSTAKA	

DAFTAR TABEL

Tabel II-1. Model <i>Confusion Matrix</i>	II-8
Tabel III-1. Rancangan Hasil <i>Confusion Matrix</i>	III-6
Tabel III-2. Rancangan Hasil Analisis Klasifikasi	III-7
Tabel III-3. Perencanaan Aktivitas Penelitian dalam bentuk WBS.....	III-10
Tabel IV-1. Kebutuhan Fungsional	IV-3
Tabel IV-2. Kebutuhan Non-Fungsional	IV-3
Tabel IV-3. Contoh Data <i>Review</i>	IV-6
Tabel IV-4. Hasil Proses <i>Case Folding</i>	IV-7
Tabel IV-5. Hasil Proses Normalisasi	IV-8
Tabel IV-6. Hasil Proses <i>Stopwords Removal</i>	IV-9
Tabel IV-7. Hasil Proses <i>Stemming</i>	IV-10
Tabel IV-8. Hasil Proses <i>Tokenization</i>	IV-11
Tabel IV-9. Hasil Perhitungan TF-IDF.....	IV-13
Tabel IV-10. Hasil Pembobotan Kata TF-IDF.....	IV-15
Tabel IV-11. Contoh Hasil Proses <i>Negation Handling</i>	IV-17
Tabel IV-12. Hasil Perhitungan TF-IDF Seluruh Data	IV-18
Tabel IV-13. Hasil Pembobotan Kata TF-IDF Seluruh Data	IV-20
Tabel IV-14. Contoh Data <i>Testing</i> Metode <i>Naïve Bayes</i>	IV-22
Tabel IV-15. Hasil Klasifikasi Data <i>Testing</i> Metode <i>Naïve Bayes</i>	IV-22
Tabel IV-16. Hasil Klasifikasi Data <i>Testing</i> Metode SVM	IV-23
Tabel IV-17. Contoh <i>Confusion Matrix</i>	IV-24
Tabel IV-18. Definisi <i>Actor</i>	IV-24
Tabel IV-19. Definisi <i>Use Case</i>	IV-26
Tabel IV-20. Skenario Melakukan <i>Text Preprocessing</i>	IV-27
Tabel IV-21. Skenario Klasifikasi <i>Naïve Bayes</i>	IV-27
Tabel IV-22. Skenario Klasifikasi <i>Support Vector Machine (SVM)</i>	IV-28
Tabel IV-23. Skenario Klasifikasi Pada <i>Input User</i>	IV-29
Tabel IV-24. Rancangan Data	IV-31
Tabel IV-25. Implementasi Kelas	IV-39
Tabel IV-26. Rencana Pengujian <i>Use Case</i> Melakukan <i>Text Preprocessing</i> ..	IV-41
Tabel IV-27. Rencana Pengujian <i>Use Case</i> Klasifikasi <i>Naïve Bayes</i>	IV-42
Tabel IV-28. Rencana Pengujian <i>Use Case</i> Klasifikasi SVM	IV-42
Tabel IV-29. Rencana Pengujian <i>Use Case</i> Klasifikasi Pada <i>Input Use</i>	IV-43
Tabel IV-30. Pengujian <i>Use Case</i> Melakukan <i>Text Preprocessing</i>	IV-43
Tabel IV-31. Pengujian <i>Use Case</i> Klasifikasi <i>Naïve Bayes</i>	IV-44
Tabel IV-32. Pengujian <i>Use Case</i> Klasifikasi SVM	IV-45
Tabel IV-33. Pengujian <i>Use Case</i> Klasifikasi Pada <i>Input User</i>	IV-46

Tabel V-1. Hasil <i>Confusion Matrix</i> Metode NB Pada Data <i>Review Spotify</i>	V-2
Tabel V-2. Hasil Evaluasi Metode <i>Naïve Bayes</i>	V-3
Tabel V-3. Hasil <i>Confusion Matrix</i> Metode SVM Pada Data <i>Review Spotify</i>	V-3
Tabel V-4. Hasil Evaluasi Metode <i>Support Vector Machine (SVM)</i>	V-4
Tabel V-5. Perbandingan Waktu Komputasi Pada Sistem Klasifikasi	V-5

DAFTAR GAMBAR

Gambar II-1. Contoh <i>Flowchart</i> Sistem Analisis Sentimen	II-1
Gambar II-2. Contoh Proses <i>Case Folding</i>	II-2
Gambar II-3. Contoh Proses <i>Tokenization</i>	II-3
Gambar II-4. Contoh Proses Normalisasi	II-3
Gambar II-5. Contoh Proses <i>Stopwords Removal</i>	II-4
Gambar II-6. Contoh Proses <i>Stemming</i>	II-4
Gambar II-7. Model <i>Rational Unified Process</i>	II-11
Gambar III-1. Diagram Kerangka Kerja	III-2
Gambar III-2. Diagram Tahapan <i>Text Preprocessing</i>	III-3
Gambar IV-1. Contoh Hasil <i>Accuracy, Precision, Recall, dan F-Measure</i> ...	IV-24
Gambar IV-2. <i>Use Case</i> Diagram	IV-25
Gambar IV-3. Rancangan Desain <i>Interface</i>	IV-32
Gambar IV-4. Rancangan Desain <i>Interface</i> Hasil Pengujian	IV-32
Gambar IV-5. Rancangan Desain <i>Interface</i> Hasil Prediksi	IV-32
Gambar IV-6. Diagram Aktivitas Melakukan <i>Text Preprocessing</i>	IV-34
Gambar IV-7. Diagram Aktivitas Klasifikasi <i>Naïve Bayes</i>	IV-34
Gambar IV-8. Diagram Aktivitas Klasifikasi SVM	IV-35
Gambar IV-9. Diagram Aktivitas Klasifikasi <i>Input User</i>	IV-35
Gambar IV-10. Diagram <i>Sequence</i> Melakukan <i>Text Preprocessing</i>	IV-36
Gambar IV-11. Diagram <i>Sequence</i> Klasifikasi <i>Naïve Bayes</i>	IV-36
Gambar IV-12. Diagram <i>Sequence</i> Klasifikasi SVM	IV-37
Gambar IV-13. Diagram <i>Sequence</i> Klasifikasi <i>Input User</i>	IV-37
Gambar IV-14. Diagram Kelas	IV-38
Gambar IV-15. <i>Interface</i> Mengunggah <i>File Dataset</i>	IV-40
Gambar IV-16. <i>Interface</i> Klasifikasi <i>Naïve Bayes</i>	IV-40
Gambar IV-17. <i>Interface</i> Klasifikasi SVM	IV-40
Gambar IV-18. <i>Interface</i> Klasifikasi Pada <i>Input User</i>	IV-40
Gambar V-1. Grafik Perbandingan Metode NB dan Metode SVM	V-6

BAB I

PENDAHULUAN

1.1 Pendahuluan

Bab pendahuluan akan membahas latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan. Bab ini juga akan memuat penjelasan mengenai gambaran umum dari keseluruhan kegiatan penelitian yang dikerjakan.

Pendahuluan dimulai dengan membahas mengenai *Sentiment Analysis* serta penelitian yang berkaitan dengan *Naïve Bayes* dan *Support Vector Machine (SVM)*.

1.2 Latar Belakang

Pada era saat ini, teknologi informasi dan komunikasi dimanfaatkan secara baik dan maksimal. Teknologi menjadi salah satu gaya hidup bagi banyak individu karena perkembangannya yang pesat khususnya pada media *streaming* musik. Pengguna dapat mendengarkan musik secara cepat dan mudah melalui media *streaming* musik. Salah satu media *streaming* musik yang memiliki banyak pengguna adalah *Spotify*.

Salah satu bidang yang ada pada *Natural Language Processing (NLP)* adalah analisis sentimen. Analisis sentimen dapat dilakukan terhadap suatu ulasan atau *review* pada aplikasi yang terdapat pada *Google Play Store*. *Google Play Store* menyediakan berbagai macam aplikasi kepada penggunanya, salah satu jenis aplikasi yang digemari di era sekarang adalah media *streaming* musik. Salah satu media *streaming* musik yang populer di dunia adalah *Spotify*. Pada saat ini, *Spotify* menjadi *streaming platform* yang paling populer di dunia karena memiliki lebih dari 300 juta pengguna. Kualitas layanan suatu aplikasi dapat dilihat melalui

kolom *review* yang ada pada *Google Play Store* (Rhajendra & Trianasari, 2021). *Review* atau ulasan merupakan hasil dari penyampaian opini, kritik, dan saran yang dapat dimanfaatkan untuk menganalisis sentimen pengguna terhadap suatu aplikasi. Hal tersebut disebabkan karena *review* mengandung sentimen yang dapat dijadikan sebagai parameter dalam meningkatkan layanan, kualitas, dan kepuasan pengguna ke depannya. Dalam menentukan sentimen dari suatu *review*, dapat dilakukan dengan mengklasifikasikannya ke dalam dua kelas, yaitu kelas positif dan kelas negatif (Yang et al., 2020). Hal ini menunjukkan bahwa analisis sentimen *review* pada suatu aplikasi tidak hanya dimanfaatkan sebagai evaluasi namun juga bermanfaat bagi pengguna lain sebagai referensi.

Algoritma *machine learning* memiliki banyak jenis, salah satunya adalah *supervised learning*. *Supervised learning* merupakan suatu teknik yang digunakan untuk mengumpulkan atau menghasilkan *output* data berdasarkan dari pengalaman atau kejadian yang sebelumnya pernah terjadi. Beberapa algoritma yang menggunakan teknik *supervised learning* antara lain yaitu *Naïve Bayes*, *Support Vector Machine*, *Decision Tree*, *Perceptron*, dan *Logistic Regression* (Ibrahim & Abdulazeez, 2021). Pada penelitian ini, algoritma yang akan digunakan yaitu *Naïve Bayes* dan *Support Vector Machine* (SVM). Algoritma *Naïve Bayes* merupakan *classifier* yang menerapkan metode probabilitas dan statistik untuk memprediksi peluang di masa depan berdasarkan pengalaman di masa lalu. Algoritma ini memiliki kinerja yang baik bahkan terhadap data yang berjumlah kecil (Villavicencio et al., 2021). Selanjutnya, algoritma *Support Vector Machine* lebih efisien dalam hal meminimalkan kesalahan prediksi dan model kompleksitas (Suasnawa et al., 2023).

Seiring berjalannya waktu terdapat permasalahan yang timbul dalam penelitian analisis sentimen yaitu dalam menentukan algoritma terbaik yang akan digunakan dari beberapa algoritma yang ada (Dang, García, & De la Prieta, 2020). Permasalahan ini dapat diatasi dengan cara melakukan perbandingan terhadap dua atau lebih metode dengan berfokus pada model dan performa yang dihasilkan oleh masing-masing metode. Dengan demikian, penelitian yang dilakukan mampu menyajikan studi perbandingan metode untuk menganalisis polarisasi sentimen terhadap data yang akan diuji (Dey et al., 2020).

Dalam penelitian (Rahat et al., 2019) hasil penelitian yang melakukan perbandingan antara metode *Naïve Bayes* dan metode *Support Vector Machine* (SVM) pada analisis sentimen terhadap data *airline reviews* memperoleh selisih hasil akurasi sebesar 5.92%. Metode *Support Vector Machine* (SVM) menghasilkan akurasi sebesar 82.48% sehingga mengungguli metode *Naïve Bayes* yang memiliki akurasi sebesar 76.56%. Hal ini berbanding terbalik dengan penelitian (Golpour et al., 2020), metode *Naïve Bayes* terbukti lebih efektif dibandingkan metode *Support Vector Machine* (SVM) dalam memprediksi variabel untuk menentukan hasil dari data *angiography test*. Hal ini disebabkan karena metode *Naïve Bayes* hanya menguji tiga variabel penting saja sehingga memiliki waktu proses yang lebih efisien dibandingkan metode *Support Vector Machine* (SVM).

Berdasarkan hal tersebut, maka penelitian ini akan melakukan pengujian terhadap metode *Naïve Bayes* dan metode *Support Vector Machine* (SVM) pada analisis sentimen data *review* aplikasi *Spotify*.

1.3 Rumusan Masalah

Berdasarkan penjelasan latar belakang tersebut, rumusan masalah dari penelitian ini adalah bagaimana kinerja yang dihasilkan oleh metode *Naïve Bayes* dan metode *Support Vector Machine* (SVM) pada analisis sentimen data *review* aplikasi *Spotify*. Berikut beberapa *research question* (RQ) dalam penelitian ini yaitu:

1. Bagaimana cara mengembangkan analisis sentimen menggunakan metode *Naïve Bayes*?
2. Bagaimana cara mengembangkan analisis sentimen menggunakan metode *Support Vector Machine* (SVM)?
3. Bagaimana kinerja yang dihasilkan oleh metode *Naïve Bayes* pada analisis sentimen data *review* aplikasi *Spotify* berdasarkan nilai *accuracy*, nilai *recall*, nilai *precision*, nilai *f-measure*, dan waktu komputasi?
4. Bagaimana kinerja yang dihasilkan oleh metode *Support Vector Machine* (SVM) pada analisis sentimen data *review* aplikasi *Spotify* berdasarkan nilai *accuracy*, nilai *recall*, nilai *precision*, nilai *f-measure*, dan waktu komputasi?
5. Bagaimana perbandingan hasil evaluasi yang dihasilkan oleh metode *Naïve Bayes* dan metode *Support Vector Machine*?

1.4 Tujuan Penelitian

Berikut tujuan dari penelitian ini yaitu:

1. Menghasilkan sistem analisis sentimen menggunakan metode *Naïve Bayes*.
2. Menghasilkan sistem analisis sentimen menggunakan metode *Support Vector Machine* (SVM).

3. Mengetahui hasil kinerja dari metode *Naïve Bayes* pada analisis sentimen data *review* aplikasi *Spotify*.
4. Mengetahui hasil kinerja dari metode *Support Vector Machine* (SVM) pada analisis sentimen data *review* aplikasi *Spotify*.
5. Mengetahui perbandingan nilai evaluasi yang dihasilkan oleh metode *Naïve Bayes* dan metode *Support Vector Machine*.

1.5 Manfaat Penelitian

Berikut manfaat dari penelitian ini yaitu:

1. Memahami mekanisme klasifikasi dari *Naïve Bayes* dan *Support Vector Machine* (SVM) pada analisis sentimen data *review* aplikasi *Spotify*.
2. Hasil penelitian dapat dimanfaatkan untuk mengetahui opini pengguna aplikasi *streaming* musik *Spotify* dan bisa dimanfaatkan sebagai referensi pada penelitian terkait.

1.6 Batasan Masalah

Berikut batasan masalah pada penelitian ini yaitu:

1. Data yang digunakan merupakan *review* atau ulasan terhadap aplikasi *Spotify* yang berasal dari *Google Play Store* dalam Bahasa Indonesia.
2. Sentimen atau *review* yang digunakan tidak mengandung emoji atau *emoticon*.
3. Data yang digunakan terdiri dari kelas positif dan kelas negatif.
4. Data yang digunakan disimpan pada file berekstensi (*.xlsx).

1.7 Sistematika Penulisan

Berikut sistematika penulisan pada penelitian ini yaitu:

BAB I. PENDAHULUAN

Bab ini berisi pembahasan mengenai latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistematika penulisan pada penelitian ini.

BAB II. KAJIAN LITERATUR

Bab ini berisi pembahasan mengenai dasar-dasar teori mengenai analisis sentimen, *text preprocessing*, metode *Naïve Bayes*, dan metode *Support Vector Machine* (SVM). Bab ini juga membahas penelitian-penelitian terdahulu yang memiliki relevansi dengan penelitian ini.

BAB III. METODOLOGI PENELITIAN

Bab ini berisi pembahasan mengenai metodologi dan tahapan perancangan penelitian seperti pengumpulan data, metode pengembangan perangkat lunak, dan manajemen proyek penelitian.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Bab ini berisi pembahasan mengenai setiap tahapan pengembangan perangkat lunak yang dikerjakan. Pengembangan sistem analisis sentimen dilakukan dengan menggunakan metode *Naïve Bayes* dan metode *Support Vector Machine* (SVM). Proses

pengembangan perangkat lunak dirancang berdasarkan model *Rational Unified Process* (RUP).

BAB V. HASIL DAN ANALISIS PENELITIAN

Bab ini berisi hasil pengujian pada perangkat lunak yang telah dirancang dan bab ini juga akan memaparkan pembahasan mengenai analisis dari hasil pengujian yang telah dikerjakan.

BAB VI. KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari hasil penelitian yang telah dikerjakan serta saran yang dapat dimanfaatkan untuk penelitian berikutnya.

1.8 Kesimpulan

Berdasarkan latar belakang penelitian yang telah diuraikan maka penelitian ini akan menguji metode *Naïve Bayes* dan metode *Support Vector Machine* (SVM) pada analisis sentimen data *review* aplikasi *Spotify*.

DAFTAR PUSTAKA

- Abbas, M., Memon, K. A., Jamali, A. A., Memon, S., & Ahmed, A. (2019). Multinomial Naive Bayes classification model for sentiment analysis. *IJCSNS 2019 - International Journal of Computer Science and Network Security*, 19(3), 62. <https://doi.org/10.13140/RG.2.2.30021.40169>
- Aldisa, R. T., & Maulana, P. (2022). Analisis Sentimen Opini Masyarakat Terhadap Vaksinasi Booster COVID-19 Dengan Perbandingan Metode Naive Bayes, Decision Tree dan SVM. *Building of Informatics, Technology and Science (BITS)*, 4(1), 106-109. <https://doi.org/10.47065/bits.v4i1.1581>
- Aninditya, A., Hasibuan, M. A., & Sutoyo, E. (2019). Text Mining Approach Using TF-IDF and Naive Bayes for Classification of Exam Questions Based on Cognitive Level of Bloom's Taxonomy. *IoTaIS 2019 - International Conference on Internet of Things and Intelligence System*, IEEE, pp. 112-117. <https://doi.org/10.1109/IoTaIS47347.2019.8980428>
- Anjali Devi, S., Sapkota, P., Rohit Kumar, K., Pooja, S., & Sandeep, M. S. (2020). Comparison of classification algorithms on twitter data using sentiment analysis. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(5), 8170-8173. <https://doi.org/10.30534/ijatcse/2020/179952020>
- Başarslan, M. S., & Kayaalp, F. (2020). Sentiment analysis with machine learning methods on social media. *Advances in Distributed Computing and Artificial Intelligence Journal (ADCAIJ)*, 9(3), 5-15. <https://doi.org/10.14201/ADCAIJ202093515>

- Brownlee, J. (2020). *Data preparation for machine learning: data cleaning, feature selection, and data transforms in Python*, Machine Learning Mastery.
- Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020). Sentiment analysis based on deep learning: A comparative study. *Electronics*, 9(3), 483. <https://doi.org/10.3390/electronics9030483>
- Darwis, D., Siskawati, N., & Abidin, Z. (2021). Penerapan Algoritma Naive Bayes Untuk Analisis Sentimen Review Data Twitter Bmkg Nasional. *Jurnal Tekno Kompak*, 15(1), 131-145. <https://doi.org/10.33365/jtk.v15i1.774>
- Dey, S., Wasif, S., Tonmoy, D. S., Sultana, S., Sarkar, J., & Dey, M. (2020). A comparative study support vector machine and Naive Bayes classifier for sentiment analysis on Amazon product reviews. *IC3A 2020 - International Conference on Contemporary Computing and Applications*. IEEE, pp. 217-220. <https://doi.org/10.1109/IC3A48958.2020.233300>
- Fitri, E. (2020). Analisis Sentimen Terhadap Aplikasi Ruangguru Menggunakan Algoritma Naive Bayes, Random Forest Dan Support Vector Machine. *Jurnal Transformatika*, 18(1), 71-80. <https://doi.org/10.26623/transformatika.v18i1.2317>
- Golpour, P., Ghayour-Mobarhan, M., Saki, A., Esmaily, H., Taghipour, A., Tajfard, M., Ghazizadeh, H., Moohebaty, M., & Ferns, G. A. (2020). Comparison of support vector machine, naïve Bayes and logistic regression for assessing the necessity for coronary angiography. *International Journal of Environmental Research and Public Health*, 17(18), 6449. <https://doi.org/10.3390/ijerph17186449>

- Ibrahim, I., & Abdulazeez, A. (2021). The role of machine learning algorithms for diagnosing diseases. *Journal of Applied Science and Technology Trends*, 2(01), 10-19. <https://doi.org/10.38094/jastt20179>
- Kristiyanti, D. A., Putri, D. A., Indrayuni, E., Nurhadi, A., & Umam, A. H. (2020). E-Wallet Sentiment Analysis Using Naïve Bayes and Support Vector Machine Algorithm. *Journal of Physics: Conference Series*, 1641(1), 012079. <https://doi.org/10.1088/1742-6596/1641/1/012079>
- Laia, M. L., & Setyawan, Y. (2020). Perbandingan hasil klasifikasi curah hujan menggunakan metode SVM dan NBC. *Jurnal Statistika Industri dan Komputasi*, 5(02), 51-61.
- Lishania, I., Goejantoro, R., & Nasution, Y. N. (2020). Perbandingan Klasifikasi Metode Naive Bayes dan Metode Decision Tree Algoritma (J48) pada Pasien Penderita Penyakit Stroke di RSUD Abdul Wahab Sjahranie Samarinda. *EKSPONENSIAL*, 10(2), 135-142.
- Nasution, M. R. A., & Hayaty, M. (2019). Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter. *J.Inform*, 6(2), 226-235.
- Novendri, R., Callista, A. S., Pratama, D. N., & Puspita, C. E. (2020). Sentiment analysis of YouTube movie trailer comments using naïve bayes. *Bulletin of Computer Science and Electrical Engineering*, 1(1), 26-32. <https://doi.org/10.25008/bcsee.v1i1.5>
- Nurkholis, A., Alita, D., & Munandar, A. (2022). Comparison of Kernel Support Vector Machine Multi-Class in PPKM Sentiment Analysis on Twitter. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 6(2), 227-233.

<https://doi.org/10.29207/resti.v6i2.3906>

Perwitasari, R., Afawani, R., & Anjarwani, S. E. (2020). Penerapan Metode Rational Unified Process (RUP) Dalam Pengembangan Sistem Informasi Medical Check Up Pada Citra Medical Centre. *Jurnal Teknologi Informasi, Komputer, dan Aplikasinya (JTIKA)*, 2(1), 76-88.

<https://doi.org/10.29303/jtika.v2i1.85>

Putranti, N. D., & Winarko, E. (2014). Analisis sentimen twitter untuk teks berbahasa Indonesia dengan maximum entropy dan support vector machine. *IJCCS 2014 - Indonesian Journal of Computing and Cybernetics Systems*.

8(1), 91-100. <https://doi.org/10.22146/ijccs.3499>

Putri, D. D., Nama, G. F., & Sulistiono, W. E. (2022). Analisis Sentimen Kinerja Dewan Perwakilan Rakyat (DPR) Pada Twitter Menggunakan Metode Naive Bayes Classifier. *Jurnal Informatika dan Teknik Elektro Terapan*, 10(1).

<https://doi.org/10.23960/jitet.v10i1.2262>

Rahat, A. M., Kahir, A., & Masum, A. K. M. (2019). Comparison of Naive Bayes and SVM Algorithm based on sentiment analysis using review dataset.

SMART 2019 - 8th International Conference System Modelling and Advancement in Research Trends. IEEE, pp. 266-270.

<https://doi.org/10.1109/SMART46866.2019.9117512>

Ramadhan, N. G., & Khoirunnisa, A. (2021). Klasifikasi Data Malaria Menggunakan Metode Support Vector Machine. *Jurnal Media Informatika*

Budidarma, 5(4), 1580-1584. <https://doi.org/10.30865/mib.v5i4.3347>

- Rana, S., & Singh, A. (2016). Comparative analysis of sentiment orientation using SVM and Naive Bayes techniques. *NGCT 2016 - 2nd International Conference on Next Generation Computing Technologies*. IEEE, pp. 106-111. <https://doi.org/10.1109/NGCT.2016.7877399>
- Sjögren, R., Stridh, K., Skotare, T., & Trygg, J. (2020). Multivariate patent analysis—Using chemometrics to analyze collections of chemical and pharmaceutical patents. *Journal of Chemometrics*, 34(1), e3041. <https://doi.org/10.1002/cem.3041>
- Suasnawa, I. W., Caturbawa, I. G. N. B., Widharma, I. G. S., Sapteka, A. A. N. G., Indrayana, I. N. E., & Sunaya, I. G. A. M. (2023). Twitter Sentiment Analysis on the Implementation of Online Learning during the Pandemic using Naive Bayes and Support Vector Machine. *iCAST-ES 2021 - 4th International Conference on Applied Science and Technology on Engineering Science*, 348-353. <https://doi.org/10.5220/0010945500003260>
- Sudhir, P., & Suresh, V. D. (2021). Comparative study of various approaches, applications and classifiers for sentiment analysis. *Global Transitions Proceedings*, 2(2), 205-211. <https://doi.org/10.1016/j.gltip.2021.08.004>
- Watrianthos, R., Giatman, M., Simatupang, W., Syafriyetti, R., & Daulay, N. K. (2022). Analisis Sentimen Pembelajaran Campuran Menggunakan Twitter Data. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 6(1), 166-170. <https://doi.org/10.30865/mib.v6i1.3383>
- Yulita, W. (2021). Analisis sentimen terhadap opini masyarakat tentang vaksin covid-19 menggunakan algoritma naïve bayes classifier. *Jurnal Data Mining dan Sistem Informasi*, 2(2), 1-9. <https://doi.org/10.33365/jdmsi.v2i2.1344>

Zhang, W., Yoshida, T., & Tang, X. (2011). A comparative study of TF* IDF, LSI and multi-words for text classification. *Expert systems with applications*, 38(3), 2758-2765. <https://doi.org/10.1016/j.eswa.2010.08.066>