

PROPOSAL SKRIPSI



**Analisis Sentimen Pengguna Aplikasi *Jobstreet* di Indonesia pada *Google Play Store* Menggunakan Metode *Naive Bayes* dan *Support Vector Machine* (SVM):
Perbandingan Akurasi Berdasarkan Rating Bintang**

Oleh:

**RANDI AFIF
NIM 101210072**

**PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
INSTITUT TEKNOLOGI DAN SAINS NAHDLATUL ULAMA
PEKALONGAN
2024**

PROPOSAL SKRIPSI

Analisis Sentimen Pengguna Aplikasi *Jobstreet* di Indonesia pada *Google Play Store* Menggunakan Metode *Naive Bayes* dan *Support Vector Machine* (SVM):
Perbandingan Akurasi Berdasarkan Rating Bintang

Oleh:

RANDI AFIF
NIM 101210072

PROGRAM STUDI TEKNOLOGI INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
INSTITUT TEKNOLOGI DAN SAINS NAHDLATUL ULAMA
PEKALONGAN
2024

HALAMAN PERSETUJUAN PEMBIMBING

Assalaamu'alaikum Wr. Wb.

Setelah saya meneliti dan mengadakan perbaikan seperlunya, bersama ini saya kirim naskah proposal skripsi saudara:

Nama : Randi Afif
NIM : 101210072
Program Studi : Teknologi Informasi
Judul : Analisis Sentimen Pengguna Aplikasi *Jobstreet* di Indonesia pada *Google Play Store* Menggunakan Metode *Naive Bayes* dan *Support Vector Machine* (SVM): Perbandingan Akurasi Berdasarkan Rating Bintang

Proposal Skripsi ini telah disetujui oleh dosen pembimbing dan siap untuk dipertahankan dihadapan Dewan Penguji proposal skripsi program Sarjana Strata 1 (S-1) Fakultas Sains dan Teknologi Institut Teknologi dan Sains Nahdlatul Ulama Pekalongan.

Demikian harap menjadikan maklum.

Wassalaamu'alaikum Wr. Wb.

Pekalongan, 15 Oktober 2024

Dosen Pembimbing,

Nur Hayati, M.Kom.
NIDN 0605109001

DAFTAR ISI

PROPOSAL SKRIPSI	i
PROPOSAL SKRIPSI	ii
HALAMAN PERSETUJUAN	iii
DAFTAR ISI	iv
DAFTAR TABEL	v
DAFTAR GAMBAR	vi
DAFTAR SINGKATAN	vii
DAFTAR SIMBOL	viii
DAFTAR LAMPIRAN	ix
BAB I PENDAHULUAN	1
A. Latar Belakang Penelitian	1
B. Rumusan Masalah	2
C. Tujuan Penelitian	3
D. Manfaat Penelitian	3
E. Batasan Penelitian	4
BAB II LANDASAN TEORI	6
A. Tinjauan Pustaka	6
1. Sentimen Analisis	6
2. Natural Language Processing (NLP)	7
2. Naive Bayes	8
3. Support Vector Machine (SVM)	9
B. Penelitian yang Relevan	10
BAB III METODE PENELITIAN	16
A. Jenis Penelitian	16
B. Sumber Data	16
C. Proses Pengumpulan Data	16
1. Web Scraping	16
2. Data Preprocessing	16
D. Metode Klasifikasi	17
1 Naive Bayes	17
2 Support Vector Machine (SVM)	17
E. Pengujian dan Evaluasi Model	17
1. Akurasi	18
2. Presisi	18
3. Recall	18
4. F1 Score	18
F. Analisis Data	18
G. Penggambaran Ulasan Positif, Negatif, dan Netral dari Tahun ke Tahun Menggunakan Regresi	18
DAFTAR PUSTAKA	20
LAMPIRAN	22

DAFTAR TABEL

Tabel 2.1. Persamaan dan Perbedaan Antara Penelitian Penulis Penelitian yang Relevan	13
---	----

DAFTAR GAMBAR

DAFTAR SINGKATAN

<i>NB</i>	: <i>Naive Bayes</i>
<i>SVM</i>	: <i>Support Vector Machine</i>
<i>KNN</i>	: <i>K-Nearest Neighbors</i>
<i>ML</i>	: <i>Machine Learning</i>
<i>AI</i>	: <i>Artificial Intelligence</i>
<i>TP</i>	: <i>True Positive</i>
<i>FP</i>	: <i>False Positive</i>
<i>TN</i>	: <i>True Negative</i>
<i>FN</i>	: <i>False Negative</i>
<i>F1</i>	: <i>F1 Score</i>
<i>R</i>	: <i>Recall</i>
<i>P</i>	: <i>Precision</i>

DAFTAR SIMBOL

$P(C X)$: Probabilitas bahwa kelas CC terjadi untuk teks XX
$P(X C)$: Probabilitas bahwa teks XX terjadi untuk kelas CC
$P(C)$: Probabilitas kelas CC
$P(X)$: Probabilitas teks XX
C	: Kelas dari sentimen (positif, negatif, netral)
X	: Teks atau fitur ulasan (kata-kata dalam ulasan)

DAFTAR LAMPIRAN

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Perkembangan teknologi digital yang pesat telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk sektor ketenagakerjaan. Salah satu perubahan yang menonjol adalah transformasi proses rekrutmen dari yang sebelumnya manual menjadi berbasis teknologi digital melalui platform online. Di Indonesia, salah satu platform pencarian kerja yang paling populer adalah Jobstreet, sebuah aplikasi yang menyediakan berbagai informasi lowongan pekerjaan dan memungkinkan perusahaan untuk merekrut talenta-talenta berkualitas. Jobstreet telah menjadi jembatan penting antara pencari kerja dan perusahaan, terutama di era digital ini.

Meskipun Jobstreet menawarkan kemudahan bagi pengguna, ulasan dan penilaian pengguna di Google Play Store menunjukkan berbagai variasi pengalaman, mulai dari kepuasan hingga ketidakpuasan. Pengguna biasanya memberikan ulasan dalam bentuk teks serta rating bintang, yang mencerminkan opini mereka tentang aplikasi. Data ini sangat penting karena dapat memberikan gambaran mengenai pengalaman pengguna, yang menjadi tolak ukur utama dalam pengembangan aplikasi lebih lanjut.

Namun, ulasan yang diberikan pengguna sering kali bervariasi dalam jumlah dan substansi, sehingga sulit untuk dianalisis secara manual. Oleh karena itu, diperlukan pendekatan otomatis yang dapat menganalisis dan mengklasifikasikan ulasan pengguna berdasarkan sentimen yang mereka sampaikan. Analisis sentimen, yang merupakan bagian dari pemrosesan bahasa alami (Natural Language Processing/NLP), menjadi solusi yang efektif untuk mengidentifikasi apakah ulasan bersifat positif, negatif, atau netral.

Dalam analisis sentimen, pemilihan algoritma yang tepat sangat penting. Penelitian sebelumnya telah menunjukkan bahwa algoritma Naïve Bayes dan SVM memiliki kelebihan masing-masing dalam klasifikasi

sentimen. Misalnya, Naïve Bayes sering kali lebih cepat dalam pelatihan dan prediksi, sementara SVM dapat memberikan hasil yang lebih baik dalam dataset yang kompleks (Wijanarto & Brilianti, 2020). Dalam penelitian ini, metode Naive Bayes dan Support Vector Machine (SVM) menjadi dua algoritma pembelajaran mesin yang akan digunakan untuk analisis sentimen. Keduanya memiliki pendekatan yang berbeda dalam mengklasifikasikan data teks, namun telah terbukti efektif dalam berbagai penelitian terkait analisis sentimen. Penelitian ini bertujuan untuk membandingkan tingkat akurasi kedua metode tersebut dalam menganalisis sentimen ulasan pengguna aplikasi Jobstreet di Indonesia, dengan mempertimbangkan rating bintang sebagai faktor pembanding utama.

Melalui analisis ini, diharapkan dapat ditemukan metode yang paling efektif untuk memahami sentimen pengguna aplikasi Jobstreet, sehingga pengembang aplikasi dapat lebih responsif dalam mengatasi masalah yang dihadapi pengguna dan meningkatkan kualitas layanan aplikasi secara keseluruhan.

B. Rumusan Masalah

Berdasarkan latar belakang yang telah terurai tersebut, peneliti membuat rumusan masalah sebagai berikut:

1. Bagaimana klasifikasi sentimen ulasan pengguna aplikasi Jobstreet di Indonesia melalui Google Play Store berdasarkan rating bintang menggunakan metode Naive Bayes dan Support Vector Machine (SVM)?
2. Bagaimana perbandingan akurasi antara metode Naive Bayes dan Support Vector Machine (SVM) dalam menganalisis sentimen ulasan pengguna aplikasi Jobstreet di Indonesia?
3. Apa faktor-faktor yang menyebabkan perbedaan akurasi antara metode Naive Bayes dan Support Vector Machine (SVM) dalam klasifikasi sentimen ulasan pengguna aplikasi Jobstreet?

C. Tujuan Penelitian

Berikut adalah tujuan penelitian dari peneliti berdasarkan rumusan masalah sebelumnya:

1. Mengklasifikasikan sentimen ulasan pengguna aplikasi Jobstreet di Indonesia melalui Google Play Store berdasarkan rating bintang menggunakan metode Naive Bayes dan Support Vector Machine (SVM).
2. Menganalisis perbandingan tingkat akurasi antara metode Naive Bayes dan Support Vector Machine (SVM) dalam menganalisis sentimen ulasan pengguna aplikasi Jobstreet di Indonesia.
3. Mengidentifikasi faktor-faktor yang mempengaruhi perbedaan akurasi antara metode Naive Bayes dan Support Vector Machine (SVM) dalam klasifikasi sentimen ulasan pengguna aplikasi Jobstreet.

D. Manfaat Penelitian

Berikut manfaat yang bisa didapat dari penelitian ini:

1. Manfaat Akademis: Penelitian ini dapat memperkaya literatur tentang penerapan metode Naive Bayes dan Support Vector Machine (SVM) dalam analisis sentimen, khususnya pada ulasan aplikasi fintech. Hasil penelitian ini dapat dijadikan referensi bagi penelitian selanjutnya yang ingin mengeksplorasi lebih jauh metode lain atau aplikasi dalam domain yang berbeda.
2. Manfaat Praktis bagi Pengembang Aplikasi: Hasil analisis sentimen ini dapat membantu pengembang aplikasi Jobstreet memahami lebih baik persepsi dan kepuasan pengguna terhadap layanan yang mereka tawarkan. Dengan memahami sentimen pengguna, pengembang dapat mengidentifikasi aspek-aspek yang perlu diperbaiki atau ditingkatkan untuk memberikan pengalaman yang lebih baik kepada pengguna.

3. Manfaat bagi Industri Fintech: Penelitian ini dapat memberikan wawasan kepada pelaku industri fintech mengenai pentingnya analisis ulasan pengguna sebagai alat evaluasi kinerja aplikasi. Dengan menggunakan pendekatan berbasis machine learning seperti Naive Bayes dan SVM, perusahaan dapat mengotomatiskan proses pemantauan ulasan, yang akan mempercepat pengambilan keputusan terkait perbaikan atau pengembangan fitur baru.
4. Manfaat bagi Pengguna Aplikasi: Dengan adanya peningkatan kualitas aplikasi yang didasarkan pada hasil analisis sentimen, pengguna aplikasi Jobstreet akan mendapatkan pengalaman yang lebih baik dalam menggunakan aplikasi, baik dari segi fungsionalitas, performa, maupun kepuasan secara keseluruhan.
5. Manfaat bagi Pengambil Keputusan: Hasil penelitian ini dapat digunakan oleh pengambil keputusan di perusahaan untuk merumuskan strategi pengembangan produk berdasarkan data yang dihasilkan dari analisis sentimen. Dengan demikian, perusahaan dapat merespons lebih cepat terhadap keluhan dan kebutuhan pengguna.

E. Batasan Penelitian

Agar penelitian ini bisa fokus dan tidak menyimpang, maka peneliti melakukan pembatasan masalah sebagai berikut:

1. Sumber Data Ulasan: Data ulasan yang digunakan dalam penelitian ini hanya berasal dari platform Google Play Store dan tidak mencakup ulasan dari platform lain seperti App Store atau forum daring lainnya.
2. Konteks Aplikasi: Penelitian ini hanya fokus pada aplikasi Jobstreet, yang merupakan aplikasi investasi, trading, dan kripto, sehingga hasil analisis sentimen mungkin tidak dapat digeneralisasi untuk aplikasi lain di kategori yang berbeda.
3. Algoritma yang Digunakan: Penelitian ini hanya membandingkan dua metode pembelajaran mesin, yaitu Naive Bayes dan Support Vector

Machine (SVM), sehingga tidak membahas metode lain yang mungkin lebih efektif dalam analisis sentimen.

4. Negara Ulasan: Penelitian ini hanya menganalisis ulasan di negara Indonesia dan mengabaikan ulasan dari negara lain yang mungkin ada di Google Play Store.
5. Akurasi Model: Tingkat akurasi yang diperoleh dari model Naive Bayes dan SVM bergantung pada teknik praproses data dan parameter yang dipilih selama pelatihan model, sehingga hasil mungkin berbeda jika parameter yang digunakan berubah.

BAB II

LANDASAN TEORI

A. Tinjauan Pustaka

1. Analisis Sentimen

Analisis sentimen merupakan proses yang penting dalam memahami opini pengguna terhadap produk dan layanan, terutama dalam konteks aplikasi belanja online. Dalam penelitian yang dilakukan oleh Subowo et al., algoritma Bidirectional Long Short Term Memory (BiLSTM) digunakan untuk menganalisis sentimen berbasis aspek pada aplikasi belanja online, seperti Cicilana. Metode ini memanfaatkan teknik pemrosesan bahasa alami (Natural Language Processing/NLP) untuk mengidentifikasi dan mengkategorikan opini yang dinyatakan dalam teks, dengan tujuan untuk menentukan sikap penulis terhadap suatu subjek, apakah positif, negatif, atau netral (Subowo et al., 2022). Dalam konteks yang lebih luas, analisis sentimen merupakan cabang dari text mining yang bertujuan untuk mengekstrak emosi dan opini dari teks. Metode SVM telah terbukti efektif dalam berbagai penelitian analisis sentimen, termasuk analisis terhadap aplikasi lain seperti Shopee dan MyPertamina, di mana penelitian-penelitian tersebut menunjukkan bahwa SVM mampu memberikan akurasi yang tinggi dalam klasifikasi sentimen (Wahyudi & Kusumawardana, 2021). Analisis sentimen berguna untuk memahami persepsi pengguna terhadap sebuah produk atau layanan secara lebih mendalam.

Beberapa teknik analisis sentimen yang digunakan secara luas termasuk:

1. Pendekatan berbasis leksikon, yang mengandalkan kata-kata yang dikaitkan dengan polaritas tertentu (positif/negatif).
2. Pendekatan berbasis pembelajaran mesin (machine learning), yang melibatkan pelatihan model untuk mengklasifikasikan sentimen berdasarkan data yang telah diberi label.

Dalam penelitian ini, pembelajaran mesin menjadi pendekatan utama untuk mengklasifikasikan sentimen ulasan aplikasi Jobstreet.

Analisis sentimen merupakan teknik yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini dalam teks, yang dapat bersifat positif, negatif, atau netral. Dalam konteks aplikasi investasi, analisis ini sangat penting karena dapat memberikan wawasan tentang kepuasan pengguna dan area yang perlu diperbaiki. Penelitian sebelumnya menunjukkan bahwa analisis sentimen dapat digunakan untuk mengevaluasi berbagai jenis produk dan layanan, termasuk aplikasi mobile, dengan akurasi yang tinggi (Adhi Putra, 2021).

2. Natural Language Processing (NLP)

Natural Language Processing (NLP) adalah cabang ilmu komputer yang berfokus pada interaksi antara komputer dan bahasa manusia. NLP memungkinkan komputer untuk memahami, menganalisis, dan menghasilkan teks dengan cara yang bermanfaat. (Syahroni, 2023) memberikan gambaran menyeluruh tentang bagaimana analisis sentimen, sebuah teknik utama dalam Natural Language Processing (NLP), dapat digunakan secara efektif dalam konteks dasbor untuk mengevaluasi umpan balik siswa dalam bahasa Indonesia. Susanto et al., (2022) memberikan gambaran menyeluruh tentang analisis sentimen dalam konteks pemrosesan bahasa alami (NLP), khususnya berfokus pada penerapan algoritma Naïve Bayes untuk mengklasifikasikan ulasan pengguna terhadap aplikasi. Dalam analisis sentimen, NLP digunakan untuk memproses ulasan teks, termasuk langkah-langkah seperti tokenisasi, penghapusan stop words, dan stemming.

Proses pra-proses data dalam NLP sangat penting untuk meningkatkan performa model pembelajaran mesin. Beberapa teknik yang sering digunakan dalam pra-proses data ulasan meliputi:

1. Tokenisasi, yaitu memecah teks menjadi unit-unit yang lebih kecil, seperti kata atau frasa.

2. Stop word removal, yaitu menghilangkan kata-kata umum yang tidak bermakna dalam konteks analisis sentimen.
3. Stemming atau lemmatization, yaitu mengubah kata-kata menjadi bentuk dasarnya untuk menyederhanakan analisis.

Teknik-teknik ini akan diterapkan dalam penelitian ini untuk memproses data ulasan pengguna aplikasi Jobstreet sebelum dilakukan analisis sentimen. Akurasi dan keandalan yang tinggi dalam mengklasifikasikan sentimen dapat dicapai dengan dukungan pengembangan NLP yang tangguh dalam konteks Indonesia (Tri Sanudin et al., 2024).

3. Naive Bayes

Naive Bayes adalah algoritma pembelajaran mesin berbasis probabilitas yang sering digunakan untuk tugas klasifikasi teks, termasuk analisis sentimen. Analisis sentimen merupakan teknik yang penting dalam memahami opini publik terhadap aplikasi, terutama dalam konteks ulasan pengguna di platform seperti Google Play Store. Dalam hal ini, algoritma Naïve Bayes telah terbukti efektif dalam mengklasifikasikan sentimen dari ulasan tersebut. Metode ini menggunakan pendekatan probabilistik untuk menentukan apakah suatu ulasan bersifat positif, negatif, atau netral, berdasarkan kata-kata yang terkandung di dalamnya (Irnawati dan Solecha, 2022). Naïve Bayes tidak hanya efektif dalam analisis sentimen, tetapi juga dalam memberikan informasi yang berguna bagi pengembang aplikasi untuk meningkatkan kualitas layanan mereka (Susanto et al., 2022).

Prinsip Dasar Naive Bayes adalah Naive Bayes bekerja dengan menghitung probabilitas bahwa teks tertentu termasuk ke dalam kelas tertentu (positif, negatif, atau netral), berdasarkan fJobstreetensi kata-kata yang ada dalam teks. Rumus dasar Naive Bayes adalah sebagai berikut:

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

Dimana:

1. $P(C|X)P(C|X)$ adalah probabilitas bahwa kelas CC terjadi untuk teks XX,
2. $P(X|C)P(X|C)$ adalah probabilitas bahwa teks XX terjadi untuk kelas CC,
3. $P(C)P(C)$ adalah probabilitas kelas CC,
4. $P(X)P(X)$ adalah probabilitas teks XX.

Dalam konteks dashboard analisis sentimen, implementasi Naïve Bayes dapat dilakukan dengan memanfaatkan data ulasan yang diambil dari Google Play Store. Data ini kemudian diproses melalui tahapan preprocessing, di mana teknik seperti TF-IDF dapat digunakan untuk mengekstrak fitur dari teks (Fibriyanti Arminda et al., 2023).

4. Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah algoritma pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. SVM bekerja dengan mencari hyperplane yang memisahkan data dalam ruang fitur ke dalam dua atau lebih kelas yang berbeda. Analisis sentimen merupakan teknik yang penting dalam memahami opini pengguna terhadap aplikasi, dan salah satu algoritma yang paling efektif untuk tujuan ini adalah Support Vector Machine (SVM). SVM telah terbukti memberikan akurasi yang tinggi dalam berbagai studi analisis sentimen, termasuk analisis terhadap aplikasi seperti Google Meet dan Pospay, di mana SVM menunjukkan kinerja yang lebih baik dibandingkan algoritma lain seperti Naïve Bayes dan K-Nearest Neighbors (Angraina dan Putri, 2022, Safryda Putri dan Ridwan, 2023, Oktafani dan Prasetyaningrum, 2022).

Dengan memanfaatkan SVM dan teknik visualisasi yang efektif, dashboard dapat memberikan informasi yang berharga bagi pengembang aplikasi untuk melakukan perbaikan yang diperlukan (Wahyu Sejati et al., 2023). Kekuatan SVM dalam analisis sentimen terletak pada kemampuannya untuk menangani data yang tidak terstruktur, seperti teks,

serta mengurangi risiko overfitting pada data yang memiliki dimensi tinggi, seperti ulasan pengguna aplikasi.

B. Penelitian yang Relevan

Ada berbagai rujukan yang diambil dalam penelitian ini guna memperkuat penelitian. Selain untuk memperkuat penelitian, rujukan-rujukan ini adalah penelitian yang sudah pernah dilakukan dan menghasilkan berbagai macam kesimpulan yang dapat dijadikan bahan perbandingan ketika akan melakukan penelitian. Tentunya mengambil penelitian ini tidak hanya perlu memperhatikan isi, namun juga apa yang telah diteliti sebelumnya oleh peneliti yang lain. Oleh karena itu, berikut beberapa penelitian terdahulu yang memperkuat penelitian ini :

1. Maulana, Voutama, dan Ridwan pada tahun 2023 menulis jurnal berjudul "Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store menggunakan Algoritma NBC" yang memberikan wawasan mendalam tentang penerapan analisis sentimen dalam konteks aplikasi mobile, khususnya aplikasi MyPertamina. Kelebihan dari jurnal ini adalah penerapan algoritma Naïve Bayes Classifier (NBC) yang sudah teruji dan terbukti efektif dalam menganalisis sentimen dari data teks, seperti ulasan pengguna di Google Play Store. Algoritma ini dipilih karena kemampuannya dalam menangani volume data yang besar dan memberikan hasil klasifikasi yang akurat. Dengan menggunakan algoritma yang teruji, penelitian ini memiliki dasar yang kuat dan memberikan keyakinan terhadap keakuratan hasil yang diperoleh. Kaitannya dengan jurnal peneliti adalah algoritma Naive Bayes Classifier (NBC) juga digunakan sebagai alat utama untuk menganalisis sentimen pengguna, dengan Support Vector Machine (SVM) sebagai algoritma pembanding. Hal ini juga menjadi inti dari penelitian yang dilakukan peneliti, di mana kedua algoritma tersebut akan diuji untuk menganalisis ulasan pengguna aplikasi Jobstreet. Sama

seperti dalam jurnal MyPertamina, penelitian ini berfokus pada ulasan pengguna di Google Play Store, yang menjadi sumber data utama untuk memahami persepsi pengguna terhadap aplikasi tersebut. Salah satu kekurangannya adalah jurnal tersebut hanya memberikan sedikit fokus pada analisis yang lebih mendalam tentang kelemahan dari Naive Bayes Classifier (NBC). Meskipun algoritma ini dikenal efektif dalam mengklasifikasikan data teks, NBC memiliki keterbatasan dalam menangani data dengan fitur yang saling bergantung, yang tidak diungkapkan secara terperinci dalam jurnal. Peneliti yang juga menggunakan NBC dalam penelitian ulasan Jobstreet mungkin perlu mempertimbangkan bagaimana kelemahan algoritma ini, seperti dalam penanganan ulasan dengan sentimen ambigu atau campuran, dapat memengaruhi hasil penelitian.

2. Syahroni dan Abd Wahab pada tahun 2023 menulis jurnal berjudul "Analisis Sentimen Komentar Mahasiswa Terhadap Dosen Mata Kuliah Pada Aplikasi SIMAT". Salah satu kelebihan dari penelitian terkait adalah penggunaan metode yang terbukti efektif dalam klasifikasi teks. Dalam kedua studi ini, algoritma Naïve Bayes digunakan sebagai salah satu pendekatan untuk menganalisis data. Naïve Bayes telah terbukti baik dalam menangani masalah klasifikasi, sehingga keduanya dapat memberikan hasil yang akurat dalam mengklasifikasikan sentimen. Keduanya juga memanfaatkan teknik analisis sentimen untuk mendapatkan wawasan dari umpan balik pengguna, baik itu komentar mahasiswa terhadap dosen atau ulasan pengguna terhadap aplikasi. Dalam hal kekurangan, analisis sentimen pada penelitian terkait juga memiliki tantangan akurasi dari analisis dapat dipengaruhi oleh kompleksitas bahasa dan konteks yang berbeda dalam komentar atau ulasan. Dalam penelitian pendidikan, komentar mahasiswa mungkin memiliki nuansa yang tidak selalu mudah diklasifikasikan, terutama jika terdapat slang atau istilah yang spesifik. Sementara itu, dalam

analisis sentimen aplikasi, ulasan pengguna juga dapat bervariasi dalam gaya penulisan dan ekspresi, yang mungkin mempersulit pengklasifikasian secara tepat.

3. Oktafani dan Prasetyaningrum pada tahun 2022 menulis jurnal berjudul “Implementasi Support Vector Machine Untuk Analisis Sentimen Komentar Aplikasi Tanda Tangan Digital” yang berisi mengenai analisis sentimen komentar pengguna aplikasi tanda tangan digital, di mana SVM terbukti efektif untuk mengklasifikasikan sentimen yang beragam menjadi positif, negatif, atau netral. Kelebihan jurnal ini menyajikan studi yang relevan dengan topik analisis sentimen yang semakin populer, terutama dalam konteks aplikasi digital yang digunakan secara luas oleh masyarakat. Fokus pada aplikasi tanda tangan digital juga menunjukkan keunikan dan relevansi topik di era digitalisasi saat ini, di mana keamanan dan efisiensi layanan menjadi perhatian utama bagi pengguna. Kaitannya dengan penelitian ini adalah sama-sama berfokus pada analisis sentimen terhadap aplikasi digital, yang menjadi semakin penting dalam konteks pengalaman pengguna. Kekurangan jurnal terkait adalah jurnal terkait menghadapi keterbatasan dalam hal jumlah dan variasi data yang digunakan. Jika dataset yang dianalisis tidak cukup besar atau tidak mewakili seluruh populasi pengguna, hasil analisis sentimen dapat menjadi tidak akurat atau bias. Hal ini berpotensi mempengaruhi generalisasi temuan terhadap aplikasi tanda tangan digital secara keseluruhan.

Berikut persamaan dan perbedaan dari penelitian ini dengan penelitian yang terdahulu:

Tabel 2.1
Persamaan dan Perbedaan Antara Penelitian Penulis dengan Penelitian yang Relevan

No	Penelitian yang Relevan	Persamaan	Perbedaan
1.	Maulana, Voutama, dan Ridwan (2023), berjudul "Analisis Sentimen Ulasan Aplikasi MyPertamina pada Google Play Store menggunakan Algoritma NBC"	<p>Metode yang Digunakan: Kedua penelitian menggunakan algoritma Naïve Bayes dan Support Vector Machine (SVM) dalam proses klasifikasi sentimen ulasan pengguna. Naïve Bayes dikenal sebagai metode yang sederhana namun efektif dalam menangani data teks, seperti ulasan di Google Play Store, sementara SVM unggul dalam kasus klasifikasi yang lebih kompleks dan biasanya memberikan akurasi yang lebih tinggi, terutama dalam data yang memiliki banyak fitur.</p> <p>Platform Google Play Store: Baik dalam penelitian jurnal terkait dengan penelitian ini, data yang dianalisis bersumber dari Google Play Store, yang memungkinkan analisis berbasis ulasan pengguna terhadap aplikasi. Platform ini</p>	<p>Konteks Aplikasi: Meskipun kedua penelitian fokus pada aplikasi mobile, aplikasi MyPertamina lebih terkait dengan layanan bahan bakar, sementara Jobstreet adalah aplikasi yang berfokus pada pencarian kerja. Hal ini mempengaruhi jenis sentimen yang diungkapkan dalam ulasan. Ulasan MyPertamina mungkin lebih banyak berfokus pada masalah teknis dan kemudahan penggunaan, sedangkan ulasan Jobstreet lebih mungkin mencakup aspek pengalaman pencarian kerja, layanan dukungan, dan informasi lowongan.</p> <p>Klasifikasi Berdasarkan Bintang: Dalam penelitian Jobstreet, klasifikasi sentimen dilakukan dengan menggabungkan analisis rating bintang dari</p>

		<p>memberikan wawasan langsung mengenai pengalaman pengguna, baik yang bersifat positif maupun negatif, serta masalah teknis yang dihadapi.</p> <p>Perbandingan Algoritma: Kedua penelitian melakukan perbandingan antara NBC dan SVM. Penelitian jurnal terkait menunjukkan efektivitas NBC dalam analisis sentimen ulasan aplikasi, namun juga membandingkannya dengan SVM untuk memberikan pandangan yang lebih komprehensif tentang metode yang paling cocok. Pada penelitian ini, metode ini akan digunakan untuk membandingkan akurasi berdasarkan rating bintang dari ulasan pengguna, yang memberikan dimensi tambahan dalam analisis.</p>	<p>ulasan pengguna. Ini memberikan perbandingan akurasi yang lebih terstruktur berdasarkan persepsi pengguna dari skor 1 hingga 5 bintang. Hal ini tidak disebutkan dalam penelitian jurnal terkait, yang lebih berfokus pada sentimen positif dan negatif secara umum.</p> <p>Hasil Sentimen: Penelitian MyPertamina menemukan bahwa sebagian besar ulasan pengguna bersifat negatif, banyak di antaranya mengeluhkan masalah teknis. Sementara itu, pada penelitian ini, sentimen mungkin lebih bervariasi karena sifat aplikasi Jobstreet yang berfokus pada karier, di mana ulasan pengguna bisa sangat beragam tergantung pada pengalaman personal mereka dalam mencari pekerjaan.</p>
2.	Syahroni dan Abd Wahab (2023) menulis jurnal berjudul "Analisis Sentimen Komentar Mahasiswa Terhadap Dosen Mata Kuliah Pada Aplikasi SIMAT"	<p>Penelitian jurnal terkait dengan penelitian ini sama-sama menggunakan algoritma Naïve Bayes digunakan sebagai salah satu pendekatan untuk menganalisis data.</p>	<p>Perbedaan dalam konteks penerapan analisis sentimen. Penelitian jurnal terkait fokus pada lingkungan pendidikan, mengamati interaksi antara mahasiswa dan dosen untuk meningkatkan kualitas pengajaran. Sebaliknya, penelitian ini lebih berorientasi pada industri dan pengalaman</p>

			pengguna terhadap layanan yang disediakan oleh aplikasi, dengan tujuan meningkatkan kepuasan pengguna. Keduanya menggunakan umpan balik sebagai dasar analisis, tetapi konteks dan fokusnya berbeda.
3.	Oktafani dan Prasetyaningrum (2022) menulis jurnal berjudul “Implementasi Support Vector Machine Untuk Analisis Sentimen Komentar Aplikasi Tanda Tangan Digital”	Kesamaan antara kedua penelitian ini adalah pentingnya pemahaman sentimen pengguna sebagai salah satu indikator utama kualitas aplikasi. Dalam aplikasi tanda tangan digital maupun aplikasi Jobstreet, pengembang dapat memanfaatkan hasil analisis sentimen untuk mengidentifikasi masalah, memperbaiki fitur aplikasi, dan merancang strategi pengembangan lebih lanjut yang sesuai dengan harapan pengguna. Oleh karena itu, meskipun kedua metode—Naive Bayes dan SVM—memiliki kekuatan dan kelemahannya masing-masing, keduanya tetap menawarkan wawasan yang berharga tentang persepsi pengguna terhadap aplikasi yang mereka gunakan.	Fokus penelitian kedua jurnal berbeda. Jurnal terkait berfokus pada analisis sentimen komentar pengguna terhadap aplikasi tanda tangan digital, sedangkan penelitian ini lebih spesifik pada analisis sentimen berdasarkan rating bintang pengguna aplikasi Jobstreet di Google Play Store. Hal ini menunjukkan bahwa konteks aplikasi yang dianalisis memiliki pengaruh pada metodologi yang digunakan dan hasil yang dicapai.

BAB III

METODOLOGI PENELITIAN

A. Jenis Penelitian

Penelitian ini merupakan penelitian kuantitatif yang menggunakan pendekatan eksperimen untuk menganalisis sentimen ulasan pengguna aplikasi Jobstreet. Fokus penelitian adalah untuk membandingkan akurasi antara dua metode pembelajaran mesin, yaitu Naive Bayes dan Support Vector Machine (SVM).

B. Sumber Data

Data yang digunakan dalam penelitian ini adalah ulasan pengguna aplikasi Jobstreet yang diambil dari Google Play Store. Ulasan ini terdiri dari teks ulasan dan rating bintang yang diberikan oleh pengguna. Pengambilan data akan dilakukan secara otomatis menggunakan web scraping dalam kurang lebih 3000 data.

C. Proses Pengumpulan Data

Proses pengumpulan data adalah langkah krusial dalam penelitian ini, yang bertujuan untuk mengumpulkan informasi yang relevan dari ulasan pengguna aplikasi Jobstreet di Google Play Store. Pengumpulan data yang efektif akan memastikan bahwa analisis sentimen yang dilakukan berdasarkan data yang akurat dan representatif. Dalam konteks ini, terdapat beberapa metode yang akan digunakan untuk mengumpulkan data ulasan, yaitu:

1. Web Scraping: Jika menggunakan teknik web scraping, peneliti akan memanfaatkan pustaka Python *google_play_scraper* untuk mengumpulkan data ulasan. Proses ini mencakup pengambilan informasi nama pengguna, rating bintang, dan isi ulasan.
2. Data Preprocessing: Setelah pengumpulan data, tahap selanjutnya adalah memproses data. Ini mencakup langkah-langkah seperti:

- a. Tokenisasi: Memecah teks ulasan menjadi kata-kata.
- b. Penghapusan Stop Words: Menghapus kata-kata umum yang tidak berkontribusi pada analisis sentimen.
- c. Stemming atau Lemmatization: Mengubah kata-kata menjadi bentuk dasar mereka untuk menyederhanakan analisis.

D. Metode Klasifikasi

Dalam penelitian ini, dua metode pembelajaran mesin yang akan digunakan dalam analisis sentimen yaitu Naive Bayes dan Support Vector Machine (SVM). Masing-masing metode memiliki karakteristik dan pendekatan yang berbeda dalam pengolahan data, sehingga akan memberikan gambaran yang komprehensif tentang bagaimana masing-masing metode dapat digunakan untuk menganalisis sentimen. Berikut penjelasannya:

1. Naive Bayes

Model ini akan dilatih menggunakan data ulasan yang telah diproses. Peneliti akan menghitung probabilitas setiap kelas (positif, negatif, netral) berdasarkan kata-kata yang ada dalam ulasan.

2. Support Vector Machine (SVM)

Model SVM juga akan dilatih menggunakan data yang sama. Peneliti akan mengatur parameter dan memilih kernel yang tepat untuk meningkatkan performa model.

E. Pengujian dan Evaluasi Model

Setelah proses pelatihan model menggunakan metode Naive Bayes dan Support Vector Machine (SVM) selesai dilakukan, langkah selanjutnya adalah melakukan pengujian dan evaluasi untuk mengukur kinerja masing-masing model. Data akan dibagi menjadi dua set, yaitu data pelatihan (training set) dan data pengujian (testing set), dengan perbandingan umum 80:20. Model akan dievaluasi menggunakan metrik berikut:

1. Akurasi: Persentase prediksi yang benar dari total prediksi.
2. Presisi: Proporsi prediksi positif yang benar dari total prediksi positif.
3. Recall: Proporsi prediksi positif yang benar dari total sebenarnya positif.
4. F1 Score: Harmonik rata-rata antara presisi dan recall.

F. Analisis Data

Hasil dari kedua model akan dianalisis dan dibandingkan. Peneliti akan menilai model mana yang memiliki akurasi tertinggi dan memberikan wawasan mengenai faktor-faktor yang mempengaruhi perbedaan akurasi antara Naive Bayes dan SVM.

G. Penggambaran Ulasan Positif, Negatif, dan Netral dari Tahun ke Tahun Menggunakan Regresi

Setelah melakukan pengujian dan evaluasi model, langkah selanjutnya adalah menggambarkan bagaimana ulasan pengguna aplikasi Jobstreet telah berubah dari waktu ke waktu. Dalam hal ini, fokus akan diberikan pada analisis tren ulasan positif, negatif, dan netral dari tahun ke tahun. Berikut penjelasannya:

1. Pengumpulan Data Historis

Untuk menggambarkan perkembangan ulasan ini, peneliti akan mengumpulkan seluruh data ulasan pengguna negara Indonesia yang telah diambil dari Google Play Store selama beberapa tahun terakhir. Data ini akan mencakup informasi tentang isi ulasan, rating bintang, serta tanggal penulisan ulasan. Dengan memiliki data yang cukup, peneliti dapat melakukan analisis tren yang lebih akurat.

2. Klasifikasi Sentimen

Untuk melakukan pengelompokan ulasan pengguna ke dalam kategori positif, negatif, dan netral, peneliti akan menggunakan algoritma

dengan akurasi tertinggi dari hasil perbandingan antara metode Naive Bayes dan SVM.

3. Analisis Regresi

Setelah ulasan diklasifikasikan, analisis regresi akan dilakukan untuk menggambarkan tren dari ulasan positif, negatif, dan netral dari tahun ke tahun. Peneliti dapat menggunakan regresi linier untuk mengidentifikasi hubungan antara waktu dan jumlah ulasan di setiap kategori. Dalam analisis ini, sumbu X dapat merepresentasikan tahun, sedangkan sumbu Y akan menunjukkan jumlah ulasan dalam masing-masing kategori.

DAFTAR PUSTAKA

- Adhi Putra, A. D. (2021). *Analisis Sentimen Pada Ulasan Pengguna Aplikasi Bibit Dan Bareksa Dengan Algoritma KNN*. 8(2), 636–646. <https://doi.org/10.35957/jatisi.v8i2.962>
- Angraina, D., & Putri, A. (2022). *Analisis Sentimen Pengguna Aplikasi Google Meet Menggunakan Algoritma Support Vector Machine*. 3(3), 472–478. <https://doi.org/10.37859/coscitech.v3i3.4260>
- Fibriyanti Arminda, N., Sulistiyowati, N., & Nur Padilah, T. (2023). *Implementasi Algoritma Multinomial Naive Bayes Pada Analisis Sentimen Terhadap Ulasan Pengguna Aplikasi Brimo*. 7(3), 1817–1822. <https://doi.org/10.36040/jati.v7i3.7012>
- Irnawati, O., & Solecha, K. (2022). *Analisis Sentimen Ulasan Aplikasi Flip Menggunakan Naive Bayes Dengan Seleksi Fitur PSO*. 4(02), 189–199. <https://doi.org/10.46772/intech.v4i02.868>
- Maulana, R., Voutama, A., & Ridwan, T. (2023). *Analisis Sentimen Ulasan Aplikasi MyPertamina Pada Google Play Store Menggunakan Algoritma NBC*. 9(1), 42–48. <https://doi.org/10.54914/jtt.v9i1.609>
- Oktafani, M., & Prasetyaningrum, P. T. (2022). *Implementasi Support Vector Machine Untuk Analisis Sentimen Komentar Aplikasi Tanda Tangan Digital*. 15(1), 10–19. <https://doi.org/10.33005/sibc.v15i1.4>
- Safryda Putri, D., & Ridwan, T. (2023). *Analisis Sentimen Ulasan Aplikasi Pospay Dengan Algoritma Support Vector Machine*. 11(01), 32–40. <https://doi.org/10.33884/jif.v11i01.6611>
- Subowo, E., Adi Artanto, F., Putri, I., & Umaedi, W. (2022). *Algoritma Bidirectional Long Short Term Memory Untuk Analisis Sentimen Berbasis Aspek Pada Aplikasi Belanja Online Dengan Cicilan*. 12(2), 132–140. <https://doi.org/10.37859/jf.v12i2.3759>
- Susanto, E. B., Paminto Agung Christianto, N., Mohammad Reza Maulana, N., & Satriedi Wahyu Binabar, N. (2022). *Analisis Kinerja Algoritma Naive Bayes Pada Dataset Sentimen Masyarakat Aplikasi NEWSAKPOLE Samsat Jawa Tengah*. 3(3), 234–241. <https://doi.org/10.37859/coscitech.v3i3.4343>

- Syahroni, A. wahab. (2023). *Analisis Sentimen Komentar Mahasiswa Terhadap Dosen Mata Kuliah Pada Aplikasi SIMAT*. 18(2). <https://doi.org/10.33998/processor.2023.18.2.1447>
- Tri Sanudin, F., Irawan, B., & Bahtiar, A. (2024). *Analisis Sentimen Ulasan Aplikasi Sapawarga Di Playstore Menggunakan Algoritma Naïve Bayes*. 8(1), 170–175. <https://doi.org/10.36040/jati.v8i1.8300>
- Wahyu Sejati, N., Ankur Singh Bist, N., & Amirsyah Tambunan, N. (2023). *Pengembangan Analisis Sentimen Dalam Rekayasa Software Engineering Menggunakan Tinjauan Literatur Sistematis*. 2(1), 95–103. <https://doi.org/10.33050/mentari.v2i1.377>
- Wahyudi, R., & Kusumawardana, G. (2021). *Analisis Sentimen Pada Aplikasi Grab Di Google Play Store Menggunakan Support Vector Machine*. 8(2), 200–207. <https://doi.org/10.31294/ji.v8i2.9681>
- Wijanarto, W., & Brilianti, S. P. (2020). *Peningkatan Performa Analisis Sentimen Dengan Resampling Dan Hyperparameter Pada Ulasan Aplikasi BNI Mobile*. 9(2), 140–153. <https://doi.org/10.30864/eksplora.v9i2.333>

LAMPIRAN