

PROPOSAL SKRIPSI



Analisis Sentimen Pengguna Aplikasi *Jobstreet* di Indonesia pada *Google Play Store* Menggunakan Algoritma *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 Algoritma *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 Algoritma *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, 25 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
3. TF-IDF	8
4. Naive Bayes	9
5. Support Vector Machine (SVM)	10
B. Penelitian yang Relevan	16
BAB III METODE PENELITIAN	19
A. Jenis Penelitian	19
B. Sumber Data	19
C. Proses Pengumpulan Data	19
1. Web Scraping	19
2. Data Preprocessing	19
D. Metode Klasifikasi	20
1 Naive Bayes	20
2 Support Vector Machine (SVM)	20
E. Pengujian dan Evaluasi Model	20
1. Prosedur Pengujian	21
2. Metode Evaluasi	21
F. Analisis Data dan kesimpulan	21
G. Visualisasi Hasil	22
H. <i>Timeline Penelitian</i>	23
I. <i>Business Modelling</i>	24
DAFTAR PUSTAKA	27
LAMPIRAN	29

DAFTAR TABEL

Tabel 2.1.	Persamaan dan Perbedaan antara Penelitian Penulis	
	Penelitian Relevan	16
Tabel 3.1	<i>Timeline</i> Penelitian	24

DAFTAR GAMBAR

Gambar 2.1	Skema Model <i>Support Vector Machines</i>	11
Gambar 3.1	Business Modelling	25

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 dan 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.

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 dilakukan analisis sentimen 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 dan negatif.

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, Algoritma 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 Algoritma Naive Bayes dan Support Vector Machine (SVM)?
2. Bagaimana perbandingan akurasi antara Algoritma 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 Algoritma 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 Algoritma Naive Bayes dan Support Vector Machine (SVM).
2. Menganalisis perbandingan tingkat akurasi antara Algoritma 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 Algoritma 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 Algoritma 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 dan negatif (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 dan negatif. 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. TF-IDF

TF-IDF (Term Frequency-Inverse Document Frequency) adalah salah satu teknik dasar dalam pengolahan teks yang digunakan untuk menilai pentingnya sebuah kata dalam dokumen tertentu, relatif terhadap sekumpulan dokumen lain. Term Frequency - Invers Document Frequency (TF-IDF) adalah metrik empiris yang menunjukkan seberapa penting suatu istilah bagi suatu dokumen dalam kumpulan dokumen (Karo Karo et al., 2023). Pada dasarnya, TF-IDF digunakan dalam information retrieval untuk membantu mesin pencari menentukan relevansi dokumen dengan kueri yang diajukan pengguna. TF-IDF mengkombinasikan dua metrik penting:

- Term Frequency (TF): Mengukur frekuensi kemunculan sebuah kata dalam dokumen. Semakin sering sebuah kata muncul dalam dokumen, semakin besar nilai TF. Rumus umum untuk TF adalah:

$$TF(t, d) = \frac{\text{Jumlah Kemunculan Kata } t \text{ dalam dokumen } d}{\text{Total Kata dalam Dokumen } d}$$

- Inverse Document Frequency (IDF): Mengukur seberapa umum atau jarang sebuah kata muncul di seluruh dokumen. Kata-kata yang sering muncul di banyak dokumen memiliki nilai IDF rendah, sementara kata yang jarang muncul di dokumen lain memiliki nilai IDF tinggi. Rumus umum untuk IDF adalah:

$$IDF(t) = \log\left(\frac{N}{DF(t)}\right)$$

di mana N adalah total jumlah dokumen, dan DF(t)DF(t) adalah jumlah dokumen yang mengandung kata t.

Perhitungan TF-IDF dilakukan dengan mengalikan kedua nilai ini:

$$TF-IDF(t, d) = Tf(t, d) \times IDF(t)$$

Hasilnya menunjukkan seberapa penting suatu kata dalam dokumen tersebut dibandingkan dengan dokumen lain. Metode TF-IDF digunakan untuk memberikan bobot pada setiap kata (Yunanda et al., 2022). Pada perhitungan bobot, TF-IDF pertama-tama menghitung nilai TF untuk setiap kata, yang menunjukkan seberapa sering kata tersebut muncul dalam dokumen tertentu. Kemudian, nilai tersebut digabungkan dengan nilai IDF yang menunjukkan seberapa jarang kata tersebut muncul di seluruh koleksi dokumen (Saputra et al., 2024).

4. 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 dan negatif, 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 dan negatif), berdasarkan frekuensi 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:

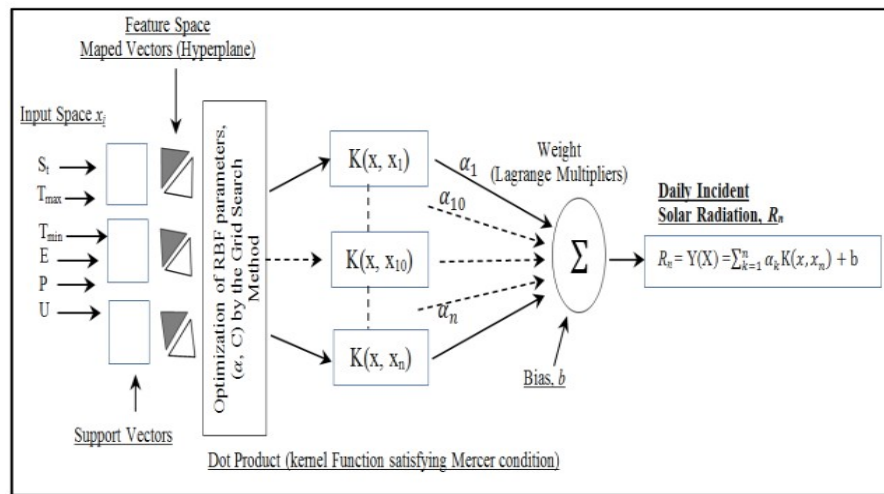
1. $P(C|X)$ adalah probabilitas bahwa kelas CC terjadi untuk teks XX,
2. $P(X|C)$ adalah probabilitas bahwa teks XX terjadi untuk kelas CC,
3. $P(C)$ adalah probabilitas kelas CC,
4. $P(X)$ adalah probabilitas teks XX.

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).

5. 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). Berikut contoh gambaran skema model SVM:



Gambar 2.1 Skema Model *Support Vector Machines*

Gambar 1 adalah gambaran skematik dari alur SVM untuk mengklasifikasikan pencahayaan matahari harian. Variabel prediktor yang digunakan sebagai input meliputi jam matahari (S_t), suhu maksimum (T_{max}), suhu minimum (T_{min}), evaporasi (E), presipitasi (P), dan kecepatan angin (U). Semua variabel ini digunakan untuk estimasi harian dari radiasi matahari global yang masuk (R_n). Dalam hal ini, SVM berusaha menemukan batas optimal (hyperplane) yang dapat memisahkan data berdasarkan variabel-variabel prediktor tersebut, sehingga dapat memberikan estimasi yang akurat untuk radiasi matahari global. Dengan cara ini, SVM dapat membantu dalam memahami bagaimana berbagai faktor meteorologis berkontribusi terhadap radiasi matahari yang diterima di permukaan bumi (Hayati et al., 2019).

Dalam konteks Support Vector Machine (SVM), terdapat beberapa jenis atau varian yang dapat digunakan tergantung pada jenis data dan masalah klasifikasi, yaitu sebagai berikut:

1. Linear SVM

- Deskripsi: Linear SVM digunakan ketika data dapat dipisahkan secara linear, yaitu jika data dapat dipisahkan oleh garis lurus (untuk kasus dua dimensi) atau hyperplane (untuk dimensi lebih tinggi).
- Cocok untuk: Data yang dapat dipisahkan dengan jelas tanpa perlu transformasi ke ruang fitur yang lebih tinggi.
- Contoh kasus: Penggunaan Linear SVM sering kali cukup pada masalah dengan dimensi data yang rendah atau ketika data dapat dipisahkan dengan baik oleh garis lurus.

Contoh penggunaan:

- Klasifikasi teks dengan jumlah fitur yang besar tetapi distribusi data linier (misalnya, analisis sentimen sederhana).

2. Non-linear SVM

- Deskripsi: Non-linear SVM digunakan ketika data tidak dapat dipisahkan secara linear. Pada kasus ini, SVM menggunakan kernel trick untuk memetakan data ke ruang fitur berdimensi lebih tinggi, di mana data tersebut dapat dipisahkan secara linear.
- Cocok untuk: Data yang kompleks dengan distribusi yang non-linear, di mana pemisahan tidak dapat dilakukan oleh hyperplane linier.
- Contoh kasus: Banyak aplikasi seperti deteksi pola pada data biologis, citra, dan klasifikasi teks multikategori memerlukan Non-linear SVM.

Jenis Kernel Non-linear yang Umum:

- Polynomial Kernel: Memperluas ruang fitur menggunakan polinomial dari derajat tertentu.
- Radial Basis Function (RBF) Kernel: Menciptakan hyperplane non-linear menggunakan fungsi Gaussian. Kernel ini sering digunakan karena fleksibilitasnya untuk menangani banyak jenis data.
- Sigmoid Kernel: Berfungsi seperti fungsi aktivasi dalam neural network.

Contoh penggunaan:

- Klasifikasi gambar atau pola wajah yang memerlukan pemisahan non-linear.

3. SVM untuk Klasifikasi Multikelas

- Deskripsi: SVM secara alami dirancang untuk masalah biner (dua kelas). Untuk menangani masalah multikelas (lebih dari dua kelas), dua pendekatan umum digunakan:
 1. One-vs-One (OvO): Membagi masalah multikelas menjadi beberapa masalah biner dengan mengambil setiap pasangan kelas, lalu membangun SVM untuk masing-masing pasangan.
 2. One-vs-All (OvA): Untuk setiap kelas, SVM dibangun untuk membedakan antara satu kelas dengan semua kelas lainnya. Hasilnya kemudian digunakan untuk mengklasifikasikan data baru.
- Cocok untuk: Klasifikasi yang melibatkan banyak kategori, seperti pengklasifikasian produk dalam berbagai kategori atau klasifikasi ulasan sebagai positif, negatif, dan netral.

Contoh penggunaan:

- Klasifikasi teks dalam beberapa kategori seperti ulasan produk dengan banyak level rating (1 sampai 5 bintang).

4. SVM untuk Regresi (Support Vector Regression - SVR)

- Deskripsi: Support Vector Regression (SVR) adalah variasi dari SVM yang digunakan untuk masalah regresi. SVR mencoba menemukan fungsi yang tidak hanya memprediksi label tetapi juga meminimalkan kesalahan dalam prediksi dengan margin tertentu.
- Cocok untuk: Masalah prediksi nilai (kontinu), seperti harga saham, suhu, atau volume penjualan.
- Contoh kasus: Prediksi suhu berdasarkan data cuaca atau prediksi harga rumah.

Contoh penggunaan:

- Prediksi harga saham atau penjualan berdasarkan fitur-fitur tertentu seperti waktu, volume transaksi, dan tren pasar.

5. Proximal SVM

- Deskripsi: Proximal Support Vector Machine (PSVM) merupakan variasi dari SVM yang mendekati solusi dengan cara meminimalkan jarak dari data ke hyperplane, bukan dengan memaksimalkan margin. PSVM biasanya lebih cepat daripada SVM standar.
- Cocok untuk: Situasi di mana pemisahan antara dua kelas tidak jelas, dan klasifikasi harus lebih fleksibel.
- Contoh kasus: Klasifikasi teks atau citra di mana margin antar kelas mungkin tidak signifikan, tetapi jarak dari hyperplane lebih dipertimbangkan.

6. Least Squares SVM (LS-SVM)

- Deskripsi: LS-SVM adalah modifikasi dari SVM yang menggunakan least squares cost function untuk menemukan hyperplane yang memisahkan data. Hal ini membuat perhitungan menjadi lebih

sederhana, dengan mengubah masalah menjadi sistem persamaan linier.

- Cocok untuk: Aplikasi di mana performa tinggi dan kecepatan diperlukan, misalnya dalam deteksi pola atau klasifikasi yang harus dilakukan secara real-time.
- Contoh kasus: Pengenalan pola wajah atau suara.

7. Structured SVM

- Deskripsi: Structured SVM digunakan untuk tugas-tugas klasifikasi yang tidak hanya menghasilkan label tunggal tetapi label yang berstruktur, seperti urutan kata, parsing bahasa alami, atau segmentasi gambar. Structured SVM memperhitungkan hubungan antar elemen dalam output yang dihasilkan.
- Cocok untuk: Masalah klasifikasi dengan output yang rumit seperti pengenalan citra atau pemrosesan bahasa alami.
- Contoh kasus: Penandaan bagian-bagian kalimat dalam Named Entity Recognition (NER) atau part-of-speech tagging dalam pemrosesan bahasa alami.

8. Transductive SVM

- Deskripsi: Transductive SVM (TSVM) digunakan untuk masalah semi-supervised learning, di mana hanya sebagian dari data yang diberi label. TSVM berusaha memprediksi label untuk data yang tidak berlabel secara bersamaan saat melatih model dengan data berlabel.
- Cocok untuk: Ketika ada data yang tidak diberi label dalam jumlah besar, dan klasifikasi masih harus dilakukan.
- Contoh kasus: Dalam analisis teks, TSVM dapat digunakan untuk membantu mengklasifikasikan teks tanpa label, berdasarkan data yang telah dilabeli.

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 ini. 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 serta persamaan dan perbedaannya:

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"	Kedua studi bertujuan untuk melakukan analisis sentimen terhadap ulasan pengguna di Google Play Store, baik pada aplikasi MyPertamina maupun Jobstreet. Keduanya juga sama-sama menggunakan algoritma klasifikasi.	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

			<p>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>
2.	<p>Syahroni dan Abd Wahab (2023) menulis jurnal berjudul "Analisis Sentimen Komentar Mahasiswa Terhadap Dosen Mata Kuliah Pada Aplikasi SIMAT"</p>	<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 pengguna terhadap layanan yang disediakan oleh aplikasi, dengan tujuan meningkatkan kepuasan pengguna. Keduanya menggunakan umpan balik sebagai dasar analisis, tetapi konteks dan fokusnya berbeda.</p>
3.	<p>Oktafani dan Prasetyaningrum (2022) menulis jurnal berjudul "Implementasi Support Vector Machine Untuk</p>	<p>Kesamaan antara kedua penelitian ini adalah pentingnya pemahaman sentimen pengguna sebagai salah satu indikator utama kualitas aplikasi. Dalam aplikasi</p>	<p>Fokus penelitian kedua jurnal berbeda. Jurnal terkait berfokus pada analisis sentimen komentar pengguna terhadap aplikasi tanda tangan digital, sedangkan</p>

	Analisis Sentimen Komentar Aplikasi Tanda Tangan Digital”	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.	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.
4.	Pratmanto, Imaniawan, dan Maarif (2023) menulis jurnal berjudul "Analisis Sentimen Pada Ulasan Pengguna Aplikasi Identitas Kependudukan Digital Dengan Metode Naive Bayes Dan K-Nearest"	Kedua penelitian menggunakan Naive Bayes sebagai salah satu metode klasifikasi utama dalam analisis sentimen. Metode ini populer dalam pengolahan teks dan analisis sentimen karena kesederhanaannya dan keefektifannya dalam banyak kasus. Fokus kedua penelitian adalah pada aplikasi analisis sentimen terhadap ulasan pengguna aplikasi digital.	Penelitian peneliti membandingkan antara Naive Bayes dan Support Vector Machine (SVM), dua metode yang sangat umum digunakan dalam klasifikasi teks. Sebaliknya, Pratmanto et al. membandingkan Naive Bayes dengan K-Nearest Neighbors (KNN), yang fokus pada klasifikasi berdasarkan kedekatan antar data.
5.	Angraina dan Putri (2022) menulis jurnal berjudul "Analisis Sentimen Pengguna Aplikasi Google Meet Menggunakan Algoritma Support Vector Machine"	Keduanya menggunakan analisis sentimen untuk mengklasifikasikan opini pengguna menjadi kategori sentimen (positif dan negatif). Hal ini menunjukkan bahwa kedua penelitian menitikberatkan pada pemahaman bagaimana pengguna merasakan aplikasi yang mereka gunakan.	Penelitian Angraina dan Putri berfokus pada aplikasi Google Meet, sementara penelitian peneliti meneliti aplikasi Jobstreet, platform pencari kerja. Fokus berbeda ini dapat mempengaruhi pola sentimen dan ulasan pengguna yang dihasilkan karena tujuan dan jenis interaksi pengguna pada kedua aplikasi tersebut tidak sama.

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. Casefolding: Menghapus tanda baca yang tidak dibutuhkan.
- b. Tokenisasi: Memecah teks ulasan menjadi kata-kata.
- c. Penghapusan Stop Words: Menghapus kata-kata umum yang tidak berkontribusi pada analisis sentimen.
- d. 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 dan 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

Pengujian dan evaluasi model adalah langkah penting dalam penelitian ini untuk memastikan bahwa model yang dibangun dapat mengklasifikasikan sentimen pengguna aplikasi Jobstreet dengan akurasi yang memadai. Dalam penelitian ini, kami menggunakan algoritma Naive Bayes sebagai metode klasifikasi dan menerapkan k-fold cross validation dengan 5 fold untuk mengevaluasi performa model.

1. Prosedur Pengujian

Pengujian dilakukan dengan membagi dataset yang berisi ulasan pengguna menjadi 5 fold. Pada setiap iterasi, satu fold digunakan sebagai data uji, sementara keempat fold sisanya digunakan untuk melatih model. Proses ini diulang sebanyak lima kali sehingga setiap fold telah digunakan sebagai data uji. Hasil dari setiap iterasi dicatat untuk analisis lebih lanjut.

2. Metode Evaluasi

Untuk mengevaluasi performa model, kami menggunakan beberapa metrik, termasuk:

- Akurasi: Proporsi prediksi yang benar dibandingkan dengan total prediksi.
- Precision: Mengukur akurasi dari prediksi positif, dihitung sebagai jumlah true positive dibagi jumlah true positive dan false positive.
- Recall: Mengukur kemampuan model untuk menangkap seluruh data positif, dihitung sebagai jumlah true positive dibagi jumlah true positive dan false negative.
- F1-Score: Merupakan rata-rata harmonis dari precision dan recall, memberikan gambaran yang lebih seimbang tentang kinerja model.

F. Analisis Data dan Kesimpulan

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. Selain itu, nantinya peneliti akan membandingkan hasilnya dengan penelitian sebelumnya agar dapat dilihat apakah hasil akurasi penelitian ini dinilai layak digunakan.

G. Visualisasi Hasil

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 dan negatif 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 dan negatif, peneliti akan menggunakan algoritma dengan akurasi tertinggi dari hasil perbandingan antara metode Naive Bayes dan SVM.

3. Visualisasi Hasil dengan *Line Chart*

Setelah ulasan diklasifikasikan, analisis regresi akan dilakukan untuk menggambarkan tren dari ulasan positif dan negatif 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.

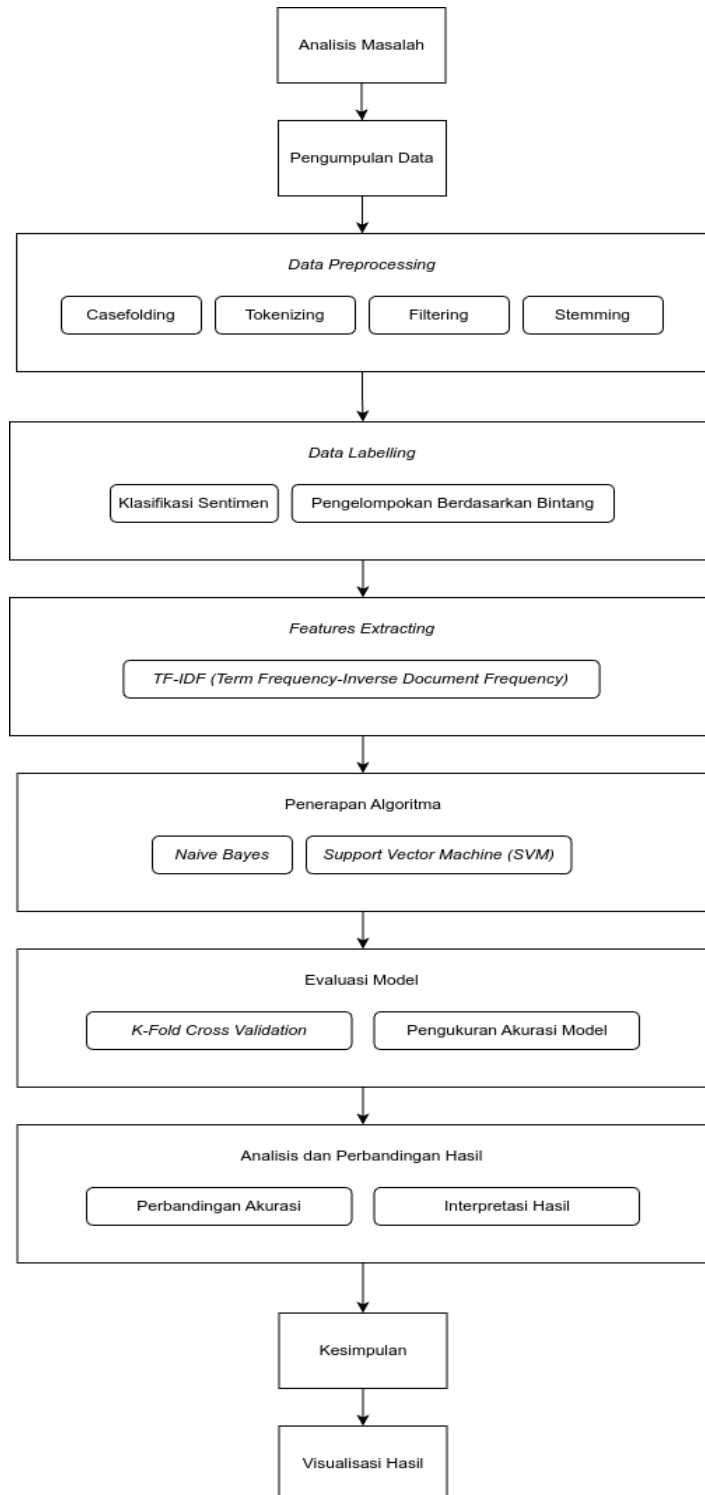
H. Timeline Penelitian

Penelitian ini dimulai dari 15 Oktober 2024 dan direncanakan selesai pada bulan Maret 2025. Rincian kegiatan penelitian dan waktu pelaksanaannya dijelaskan dalam tabel berikut:

Tabel 3.1
Timeline Penelitian

No	Kegiatan	Bulan ke:					
		1	2	3	4	5	6
1	Pencarian dan penetapan penelitian						
2	Pengumpulan Data						
3	<i>Data Preprocessing</i>						
4	<i>Data Labelling</i>						
5	<i>Features Extracting</i>						
6	Penerapan Algoritma						
7	Evaluasi Model						
8	Analisis dan Perbandingan Hasil						
9	Kesimpulan						
10	Visualisasi Hasil						

I. Tahapan Penelitian



Gambar 3.1 Tahapan Penelitian

Tahapan Penelitian dari penelitian ini adalah sebagai berikut:

1. Analisis Masalah

Pada tahap awal, peneliti menemukan bahwa peringkat bintang di Google Play Store tidak cukup menggambarkan sentimen orang Indonesia terhadap aplikasi Jobstreet, karena ulasannya bersifat global dan peringkat netral (bintang 3) menambah kebingungan. Oleh karena itu, diperlukan analisis sentimen yang lebih mendalam dengan menghapus nilai netral, sehingga dapat memfokuskan pada sentimen positif dan negatif untuk memahami kepuasan atau ketidakpuasan pengguna Indonesia dengan lebih akurat.

2. Pengumpulan Data

- Data Ulasan: Mengumpulkan ulasan pengguna aplikasi Jobstreet dari Google Play Store. Data ulasan meliputi teks ulasan, rating bintang, dan tanggal ulasan.

3. Data Preprocessing

Tahap membersihkan data menggunakan teknik casefolding, tokenizing, filtering, stemming.

4. Data Labelling

- Klasifikasi Sentimen: Setiap ulasan akan dikategorikan berdasarkan sentimen positif dan negatif dengan mempertimbangkan isi teks dan rating bintang.
- Pengelompokan Berdasarkan Bintang: Data dibagi ke dalam kelompok berdasarkan rating bintang untuk perbandingan akurasi antara berbagai kelompok.

5. Features Extracting

- Ekstraksi Fitur: Menggunakan teknik seperti TF-IDF (Term Frequency-Inverse Document Frequency) atau bag of words untuk mengubah teks menjadi representasi numerik yang dapat digunakan oleh algoritma pembelajaran mesin.

6. Penerapan Algoritma

- Naive Bayes: Menerapkan algoritma Naive Bayes untuk klasifikasi sentimen ulasan.
- Support Vector Machine (SVM): Menerapkan algoritma SVM untuk klasifikasi sentimen yang sama.

7. Evaluasi Model

- Cross Validation: Menggunakan k-fold cross validation untuk mengevaluasi kinerja masing-masing model.
- Mengukur performa model menggunakan metrik seperti akurasi.

8. Analisis dan Perbandingan Hasil

- Perbandingan Akurasi: Menganalisis perbandingan akurasi antara algoritma Naive Bayes dan SVM pada setiap kelompok data berdasarkan rating bintang.
- Interpretasi Hasil: Menjelaskan bagaimana masing-masing algoritma bekerja pada ulasan dengan rating yang berbeda, dan mengapa ada perbedaan akurasi di antara keduanya.

9. Kesimpulan

Menyimpulkan algoritma mana yang memberikan hasil terbaik untuk analisis sentimen pengguna aplikasi Jobstreet berdasarkan rating bintang.

10. Visualisasi Hasil

Visualisasi hasil menggunakan Diagram Garis (Line Chart). Diagram garis dapat digunakan untuk menunjukkan tren akurasi sepanjang setiap hasil.

Tahapan Penelitian ini dirancang untuk memastikan setiap langkah dari pengumpulan data hingga evaluasi model dilakukan secara sistematis dan menghasilkan analisis yang dapat dipertanggungjawabkan.

DAFTAR PUSTAKA

- 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>
- Hayati, N., Suryono, S., & Widodo, C. E. (2019). Fog computing uses Radial Basis Function for Power Production Classification Solar Panel Electricity. *Proceedings of 2019 4th International Conference on Informatics and Computing, ICIC 2019*, 14–19. <https://doi.org/10.1109/ICIC47613.2019.8985702>
- Karo Karo, I. M., Romia, R., Dewi, S., & Fadilah, P. M. (2023). Hoax Detection on Indonesian Tweets using Naïve Bayes Classifier with TF-IDF. *Journal of Information System Research (JOSH)*, 4(3), 914–919. <https://doi.org/10.47065/josh.v4i3.3317>
- 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>
- Saputra, N. A., Aeni, K., & Saraswati, N. M. (2024). Indonesian Hate Speech Text Classification Using Improved K-Nearest Neighbor with TF-IDF-ICSpF. *Scientific Journal of Informatics*, 11(1), 21–30. <https://doi.org/10.15294/sji.v11i1.48085>
- Susanto, E. B., Paminto Agung Christianto, N., Mohammad Reza Maulana, N., & Satriedi Wahyu Binabar, N. (2022). *Analisis Kinerja Algoritma Naïve 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>
- 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>
- Yunanda, G., Nurjanah, D., & Meliana, S. (2022). Recommendation System from Microsoft News Data using TF-IDF and Cosine Similarity Methods. *Building of Informatics, Technology and Science (BITS)*, 4(1), 277–284.
<https://doi.org/10.47065/bits.v4i1.1670>

LAMPIRAN