

## Perbandingan Evaluasi Kernel Support Vector Machine dalam Analisis Sentimen *Chatbot* AI pada Ulasan Google Play Store

Celine Mutiara Putri<sup>1</sup>, M. Afdal<sup>2</sup>, Rice Novita<sup>3</sup>, Mustakim<sup>4</sup>

Sistem Informasi, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia, 28293  
e-mail: <sup>1</sup>12050327108@students.uin-suska.ac.id, <sup>2</sup>m.afdal@uin-suska.ac.id, <sup>3</sup>rice.novita@uin-suska.ac.id, <sup>4</sup>mustakim@uin-suska.ac.id,

Submitted Date: June 27<sup>th</sup>, 2024  
Revised Date: July 17<sup>th</sup>, 2024

Reviewed Date: July 04<sup>th</sup>, 2024  
Accepted Date: July 24<sup>th</sup>, 2024

### Abstract

AI (Artificial Intelligence) is becoming very important these days due to its ability as a personal assistant to increase efficiency, automate routine tasks, and speed up manual processes. AI *chatbot* are one of the practical applications of AI in language understanding, have various benefits and drawbacks that cause various comments from users in the review column on the Google Play Store. This research discusses sentiment analysis of AI *chatbot* application reviews using four SVM kernels. Labeling uses InSet Lexicon and hyperparameters to produce the best parameters. The purpose of the research is to find out how users respond to interactions with ChatGPT, Perplexity AI, and Bing Chat and prove whether the kernel in SVM can increase the accuracy value. The percentage division between test data and training data is 70:30, 80:20, and 90:10, data labeling using 2 sentiment classes and 3 sentiment classes, and using and not using the SMOTE Oversampling technique. The experimental results obtained the highest accuracy using SVM kernel Linear scenario 90:10 with an accuracy value of 92.68%.

Keywords: ChatGPT; Perplexity AI; Bing Chat; Sentiment analysis; SVM ; Kernel; Hyperparameter

### Abstrak

Artificial Intelligence (AI) menjadi sangat penting dimasa ini karena kemampuannya sebagai asisten pribadi untuk meningkatkan efisiensi, mengotomatisasi tugas-tugas rutin, dan mempercepat proses manual. *Chatbot* AI menjadi salah satu penerapan praktis AI dalam pemahaman bahasa, memiliki berbagai manfaat serta kekurangan yang menimbulkan berbagai komentar para penggunanya pada kolom ulasan di Google Play Store. Penelitian ini membahas tentang analisis sentimen ulasan aplikasi *chatbot* AI menggunakan empat kernel SVM. Pelabelan menggunakan InSet Lexicon dan hyperparameter untuk menghasilkan parameter terbaik. Tujuan penelitian yaitu untuk mengetahui bagaimana pengguna merespons interaksi dengan ChatGPT, Perplexity AI, dan Bing Chat serta membuktikan penggunaan kernel yang tepat pada SVM dapat meningkatkan nilai akurasi. Diskenariokan persentase pembagian antara data uji dan data latih adalah 70:30, 80:20, dan 90:10, pelabelan data menggunakan 2 kelas sentimen dan 3 kelas sentimen, serta menggunakan dan tidak menggunakan teknik SMOTE Oversampling. Hasil eksperimen diperoleh akurasi tertinggi menggunakan SVM kernel Linier skenario 90:10 dengan nilai accuracy 92.68%.

Keywords: ChatGPT; Perplexity AI ; Bing Chat; Sentiment analysis; SVM; Kernel; Hyperparameter

### 1 Pendahuluan

Kemunculan Artificial Intelligence (AI) atau kecerdasan buatan menjadi sangat penting dimasa ini karena kemampuannya sebagai asisten pribadi, dan tidak mengherankan jika AI telah

memikat imajinasi banyak orang (Caramiaux & Fdili Alaoui, 2022). AI adalah bidang ilmu yang bertujuan untuk memungkinkan mesin melakukan tugas-tugas yang biasanya memerlukan kecerdasan manusia (Jarrahi, 2018). Tujuan utama



pengembangan AI bukanlah untuk menggantikan manusia dan mengambil posisi pekerjaan mereka, namun untuk meningkatkan efisiensi, mengotomatisasi tugas-tugas rutin, dan mempercepat proses manual (Zirar et al., 2023). AI memiliki cakupan aplikasi yang luas, melampaui tugas-tugas spesifik ke berbagai bidang seperti pemecahan masalah, pemodelan, permainan, robot, sistem pembelajaran dan adaptif, pemahaman visual dan bahasa (Jiang et al., 2022). AI dapat secara signifikan meningkatkan proses pengumpulan, analisis, prediksi, dan evaluasi data, menawarkan wawasan berharga kepada peneliti dan meningkatkan prosedur pengambilan keputusan (Saeidnia et al., 2024).

Bidang AI yang terus berkembang dan meningkat, menjadikannya lebih cerdas dan lebih bermanfaat (Xu et al., 2021). Salah satu penerapan praktis AI dalam pemahaman bahasa adalah pengembangan *chatbot* AI yang mendapatkan popularitas menjelang akhir November 2022 dengan diperkenalkannya ChatGPT. Dapat menghasilkan respon yang sangat mirip dengan manusia serta dapat menangani urutan inputan dengan panjang bervariasi menjadi keunggulan yang dimiliki oleh ChatGPT (Iorliam & Ingio, 2024).

Platform *chatbot* AI lain yang juga populer adalah Perplexity AI. Perplexity AI menggabungkan model bahasa yang sangat besar dengan kemampuan GPT-3. Perplexity AI mencari informasi terkait di Internet menggunakan pemrosesan bahasa alami (NLP) dan pembelajaran mesin untuk memberikan jawaban atas permintaan pengguna (Kavak, 2023). Model ini dibuat mirip dengan ChatGPT dan menentukan kemungkinan bahwa kata saat ini akan diprediksi dengan benar berdasarkan konteks kata-kata sebelumnya (Iorliam & Ingio, 2024). Perplexity AI mencantumkan sumber-sumber yang digunakannya untuk memberikan jawaban atas pertanyaan pengguna dan respon yang dihasilkan sangat ringkas dan tidak berbelit-belit (Iorliam & Ingio, 2024).

Microsoft Edge juga memiliki fitur AI *chatbot* yang memungkinkan pengguna berinteraksi untuk mendapatkan jawaban atas pertanyaan, membuat konten, dan berbagai hal lainnya yang disebut Bing Chat (Iorliam & Ingio, 2024). Bing Chat didukung oleh kecerdasan buatan, sehingga dapat belajar dari interaksi

dengan penggunaannya dan meningkatkannya dari waktu ke waktu, respon yang dihasilkan memiliki tingkat akurasi yang tinggi, serta dapat menghasilkan gambar pada kotak obrolan tempat teks dihasilkan (Rudolph et al., 2023). *Chatbot* Microsoft ini dibangun di atas arsitektur GPT-4 OpenAI dan dikembangkan dengan teknologi eksklusif yang disebut Prometheus (Ribas, 2023).

Dalam era di mana penggunaan aplikasi mobile semakin meluas, ulasan pengguna di platform seperti Google Play Store menjadi sumber informasi yang sangat berharga (Wahyu Handani et al., 2019). Ulasan tersebut tidak hanya mencerminkan pengalaman pengguna terhadap sebuah aplikasi, tetapi juga mencakup sentimen yang dapat mempengaruhi persepsi dan penerimaan publik terhadap produk tersebut (Gezici et al., 2019). Ulasan yang terus meningkat seiring berjalannya waktu dapat dimanfaatkan untuk mengetahui sentimen pengguna dengan cara melakukan analisis terhadap ulasan tersebut.

Salah satu pendekatan yang telah terbukti efektif dalam menganalisis sentimen adalah Support Vector Machine (SVM) dengan berbagai jenis kernel (Nurkholis et al., 2022). Dalam penelitian (Prastyo et al., 2020) disebutkan bahwa fungsi kernel pada SVM memainkan peran penting dalam meningkatkan kinerja SVM. Penelitian oleh (Sengkey et al., 2020) menemukan bahwa penggunaan kernel yang berbeda mempengaruhi akurasi pengklasifikasian SVM. Dalam penelitian Indonesian Sentiment Analysis: An Experimental Study of Four Kernel Functions on SVM Algorithm with TF-IDF (Prastyo et al., 2020) fungsi kernel RBF SVM+TF-IDF mencapai kinerja terbaik dengan nilai akurasi 96.61%. Pada penelitian serupa yang dilakukan oleh (Rianti et al., 2021) hasil evaluasi model klasifikasi terbaik dari penelitian tersebut adalah kernel sigmoid dengan nilai akurasi 92% serta parameter  $C=100$ ,  $\gamma=0.01$ , dan  $r=1$ . Penerapan SMOTE oversampling dan pembobotan kata TF-IDF dalam kernel polynomial berhasil menaikkan kinerja algoritma SVM sebesar 2% dalam penelitian (Rabbani et al., 2023). Sementara itu kernel linear dan sigmoid mendapatkan nilai akurasi lebih tinggi yaitu 87% dibandingkan RBF dan polynomial yang hanya 86% dalam penelitian (Permata Aulia et al., 2021).

Pada penelitian ini akan melakukan analisis sentimen review pada aplikasi ChatGPT,

Perplexity AI, dan Microsoft Bing Chat yang dikumpulkan dari Google Play Store menggunakan algoritma SVM. Kamus InSet Lexicon digunakan untuk mengklasifikasikan sentiment tersebut termasuk dalam kategori positif, negatif, atau netral. Dalam penelitian ini juga menggunakan hyperparameter untuk mengoptimalkan parameter agar mendapatkan kinerja terbaik (Auliaddina & Arifin, 2024). Penelitian ini juga mencoba mengatasi ketidak seimbangan label dengan menerapkan teknik Synthetic Minority Over-sampling Technique (SMOTE) pada dataset. SMOTE menghasilkan sampel sintetik dari kelas minoritas dengan melakukan oversampling pada setiap titik data dengan mempertimbangkan kombinasi linier dari tetangga kelas minoritas yang ada (Kumar & Prusty, 2020).

## 2 Landasan Teori

### 2.1 Analisis Sentimen

Analisis sentimen adalah disiplin ilmu yang memeriksa dan mengevaluasi pendapat atau pandangan, mengumpulkan dan memisahkan opini, menganalisis pengaruh dan emosi public

terhadap suatu peristiwa atau topik (Liu, 2022). Analisis sentimen ini membantu organisasi, bisnis, bahkan individu untuk memahami bagaimana pandangan, perasaan atau sikap masyarakat dalam merespon merespons berbagai hal yang sedang terjadi (Sari & Suryono, 2024).

### 2.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) merupakan algoritma klasifikasi yang akan memprediksi data ke daerah kelas mana yang merupakan tempat dari data tersebut (Ananda & Suryono, 2024). Model SVM membangun bidang pemisah (hyperplane) yang memiliki batas maksimum antara dua kelas, sehingga dapat memaksimalkan pemisahan antar kelas. Untuk menghasilkan hyperplane yang memisahkan kelas dengan baik dan memiliki generalisasi yang baik terhadap data, perlu dilakukan optimasi pada hyperparameter. Adapun parameter yang digunakan yaitu  $C$  (cost),  $\gamma$  (gamma),  $r$  (coefficient), dan  $d$  (degree). Penentuan parameter optimasi pada masing-masing kernel ditunjukkan pada Tabel 1.

Tabel 1. Parameter Optimasi

Kernel	Parameter Optimasi	Nlai
Linear	$C$	$C = 0.001, 0.01, 0.1, 1, 10, 100$
RBF	$C$ dan $\gamma$	$\gamma = 0.001, 0.01, 0.1, 1, 10, 100$
Sigmoid	$C, \gamma$ dan $r$	$r = 0, 1, 5, 8, 10$
Polynomial	$C, \gamma, r$ , dan $d$	$d = 0, 1, 2, 3$

Penentuan nilai parameter tersebut didasarkan pada beberapa penelitian sebelumnya yaitu penelitian (Tuttle et al., 2020) menggunakan parameter  $C = \{0.001, 0.01, 0.1, 1, 10, 100\}$ ,  $\gamma = \{0.001, 0.01, 0.1, 1, 10, 100\}$ ,  $r = \{0, 1, 5, 8, 10\}$ , dan  $d = \{2, 3, 4, 5, 6, 7, 8\}$ . Kemudian penelitian (Rudd, 2018) menggunakan parameter  $C = \{0.001, 0.01, 0.1, 1, 10\}$ ,  $\gamma = \{0.001, 0.01, 0.1, 1\}$ , dan  $d = \{2, 3\}$ . Sedangkan penelitian (Nanda et al., 2018) parameter yang digunakan adalah  $C = \{2^{-15}, 2^{-14}, \dots, 22\}$ ,  $r = \{2-10, 2-9, \dots, 22\}$ ,  $\gamma = \{2-10, 2-9, \dots, 22\}$ , dan  $d = \{0, 1, 2, 3\}$ . Maka untuk penelitian ini telah dilakukan beberapa penyesuaian dan memilih nilai yang memungkinkan untuk digunakan.

Hasil akhir klasifikasi pada algoritma SVM juga dipengaruhi oleh pemilihan kernel (Rianti et al., 2021). Fungsi kernel pada algoritma SVM

memainkan peran penting dalam meningkatkan kinerja algoritma SVM (Prastyo et al., 2020b). Empat kernel yang umum digunakan pada algoritma SVM yaitu:

- 1) Kernel Linear

$$K = (X_i, X_j) = (X_i, X_j)$$

- 2) Kernel Polynomial

$$K = (X_i, X_j) = (X_i, X_j + 1)^p$$

- 3) Kernel RBF

$$K = (X_i, X_j) = e^{-\gamma(X_i - X_j)^2}$$

- 4) Kernel Sigmoid

$$K(X_i, X_j) = \tanh(\eta X_i \cdot X_j + v)$$

### 2.3 ChatGPT

ChatGPT (Generative Pre-trained Transformer) adalah model bahasa kecerdasan buatan yang dikembangkan oleh OpenAI (Wu et



al., 2023). Ini adalah varian dari model GPT yang dimaksudkan untuk berinteraksi dengan pengguna dalam bentuk chat atau percakapan. ChatGPT menghasilkan respons yang realistis untuk setiap inputan dari pengguna sesuai dengan konteks dalam berbagai jenis percakapan (Iorliam & Ingio, 2024), tidak hanya dalam bahasa Inggris, tetapi juga bahasa lainnya dalam tingkat yang bervariasi (Ernawati & Wati, 2024).

## 2.4 Perplexity.AI

Dikenal karena fokusnya pada akurasi dan mengutip sumber yang dapat diandalkan. Aplikasi ini juga menawarkan "mode co-pilot GPT-4" yang unik untuk penulisan kolaboratif. Memiliki sekitar 2 juta pengguna aktif bulanan per April 2023 (Sullivan, 2023). Perplexity menerapkan kecerdasan buatan, pemrosesan bahasa alami, dan pembelajaran mesin untuk menawarkan pencarian Internet yang cerdas untuk menghasilkan informasi yang konvergen, akurat, dan terkini yang ditemukan di Internet dengan sumbernya masing-masing (Iorliam & Ingio, 2024).

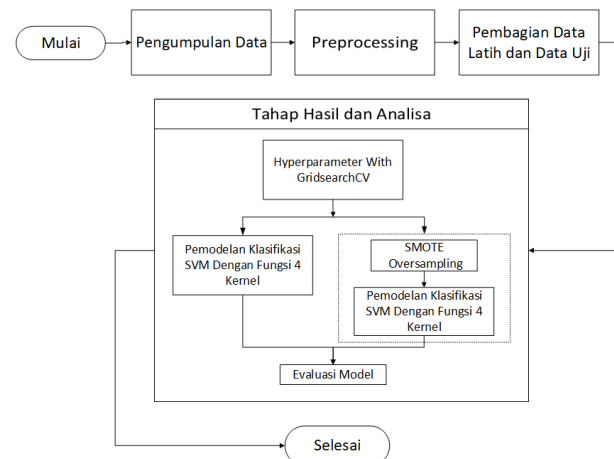
## 2.5 Bing Chat

Bing Chat dibangun sebagai perpanjangan dari mesin pencari Bing. Didukung oleh GPT-4, aplikasi ini menawarkan fitur-fitur canggih seperti dukungan multimodal, input/output visual, dan riwayat obrolan. Integrasinya dengan Bing Search menjadikannya alat yang unggul untuk mengumpulkan informasi dan interaksi percakapan, dengan lebih dari 100 juta pengguna per Maret 2023 (Iorliam & Ingio, 2024).

## 3 Metode Penelitian

Penelitian ini terdiri dari serangkaian proses yang terencana dan sistematis untuk menemukan jawaban dari permasalahan yang diangkat. Berikut ini adalah tahapan-tahapan yang dilakukan dalam penelitian yang dapat dilihat pada Gambar 1.

Dalam penelitian ini dilakukan perbandingan evaluasi kernel SVM dalam analisis sentiment chatbot AI. Adapun data yang digunakan berupa data ulasan dari aplikasi ChatGPT, Perplexity, dan Bing Chat yang diambil dari Google Play Store melalui scraping data menggunakan Library Google-Play-Scraper menggunakan bahasa pemrograman python.



Gambar 1. Metodologi Penelitian

Data yang terkumpul dari proses scraping masih dalam bentuk unstructured data, sehingga perlu dilakukan preprocessing agar data menjadi lebih terstruktur. Proses preprocessing yang dilakukan dalam penelitian ini melibatkan beberapa tahapan, yaitu pembersihan data dari emoji dan karakter yang tidak diperlukan (cleaning text), pengubahan huruf menjadi huruf kecil (case folding), penyaringan kata (filtering), pemisahan kata menjadi token (tokenizing), menghapus konjungsi (stopword removal), mengubah kebentuk kata dasar (stemming), pelabelan kata, pembobotan kata (TF-IDF), dan setelahnya dilakukan pelabelan data dengan 2 kelas sentimen yakni ulasan positif, ulasan negatif, dan 3 kelas sentiment yakni ulasan positif, ulasan negatif dan ulasan netral.

Data yang telah melalui tahapan preprocessing selanjutnya akan dibagi menjadi data latih dan data uji. Dalam penelitian ini, akan dicobakan tiga skenario pembagian data latih dan data uji yaitu : 70% data latih dan 30% data uji, 80% data latih dan 20% data uji, dan 90% data latih dan 10% data uji.

Tahap selanjutnya yaitu proses pencarian grid (gridsearch) terhadap berbagai kombinasi hyperparameter untuk menemukan kombinasi yang akan menghasilkan model yang berfungsi dengan baik pada setiap kernel menggunakan teknik GridsearchCV menggunakan library scikit-learn. Dilanjutkan dengan pemodelan menggunakan algoritma SVM dengan empat kernel yang kemudian akan diuji performanya menggunakan confusion matrix. Teknik SMOTE oversampling juga akan dilakukan sebagai variasi pengujian terhadap data. Dari confusion matrix

akan diperoleh akurasi dari masing-masing pengujian kernel dan selanjutnya masuk ke tahap visualisasi dengan menggunakan word cloud.

## 4 Hasil dan Pembahasan

### 4.1 Pengumpulan Data

Data awal yang digunakan dalam penelitian ini adalah ulasan dari aplikasi ChatGPT, Perplexity, dan Bing Chat pada Google Play Store yang dikumpulkan melalui scraping data menggunakan Library Google-Play-Scraper menggunakan bahasa pemrograman python sebanyak 1000 data.

### 4.2 Preprocessing Data

Sebanyak 1000 data ulasan yang didapat saat scrapnging akan melewati tahap preprocessing. Preprocessing dilakukan untuk membersihkan dan mengubah teks mentah menjadi format yang lebih mudah untuk diproses oleh algoritma pembelajaran mesin dan mudah untuk dianalisa. Proses yang dilakukan meliputi pembersihan data dari emoji dan karakter yang tidak diperlukan (cleaning text), pengubahan huruf menjadi huruf kecil (case folding), penyaringan kata (filtering), pemisahan kata menjadi token (tokenizing), menghapus konjungsi (stopword removal), mengubah kebentuk kata dasar (stemming), dan pembobotan kata(TF-IDF)

Tabel 2. Hasil Preprocessing Data

Ulasan	Cleaning	Tokenizing	Filtering	Steming
Aplikasinya sangat bagus dan keren👍😊, mungkin kalau bisa ditambahkan fitur mencari teks soalnya suka lupa apa yang sudah di tanya jadinya nanya lagi, biar enggak dua kali kerjaan.	aplikasinya sangat bagus dan keren mungkin kalau bisa ditambahkan fitur mencari teks soalnya suka lupa apa yang sudah di tanya jadinya nanya lagi biar enggak dua kali kerjaan.	['aplikasinya', 'sangat', 'bagus', 'dan', 'keren', 'mungkin', 'kalau', 'bisa', 'ditambahkan', 'fitur', 'mencari', 'teks', 'soalnya', 'suka', 'lupa', 'apa', 'yang', 'sudah', 'di', 'tanya', 'jadinya', 'nanya', 'lagi', 'biar', 'enggak', 'dua', 'kali', 'kerjaan']	['aplikasinya', 'bagus', 'keren', 'fitur', 'mencari', 'teks', 'suka', 'lupa', 'nanya', 'biar', 'enggak', 'kali', 'kerjaan']	['aplikasi', 'bagus', 'keren', 'fitur', 'cari', 'teks', 'suka', 'lupa', 'nanya', 'biar', 'enggak', 'kali', 'kerja']

Setelah dilakukannya *preprocessing*, data yang telah dibersihkan akan dihitung label sentimennya untuk mengetahui kecenderungan pengguna terhadap ketiga Aplikasi *chatbot* AI. Proses pelabelan dilakukan menggunakan kamus InSet Lexicon. Pada penelitian ini data akan dilabelakan menjadi 3 kelas sentimen yaitu positif, negatif dan netral, dan juga 2 kelas sentimen yaitu positif, dan negatif. Hasil pelabeian disajikan pada Tabel 3 dan Tabel 4

Tabel 3. Rincian hasil pelabelan 2 kelas sentimen

Jumlah	2 Kelas	
	Tidak Seimbang	Seimbang
Positif	614	614
Negatif	338	614

Tabel 4. Rincian hasil pelabelan 3 kelas sentimen

Jumlah	3 Kelas	
	Tidak Seimbang	Seimbang
Positif	493	493

Negatif	338	493
Netral	121	493

### 4.3 Pembagian Data Latih dan Data Uji

Data yang telah melalui proses preprocessing akan dibagi menjadi data latih dan data uji. Penerapan pembagian data latih dan data uji dalam penelitian ini akan dibagi menjadi tiga yaitu 70:30, 80:20, dan 90:10. Serta penerapan teknik SMOTE *oversampling* dan tanpa penggunaan teknik SMOTE *oversampling* pada data sebagai variasi perbandingan.

### 4.4 Hasil dan Analisa

Proses selanjutnya adalah mengimplementasikan data kedalam algoritma SVM dan mencari kombinasi optimal dari nilai-nilai hyperparameter untuk meningkatkan kinerja model dengan empat kernelnya yaitu, Linier, RBF, Polynomial, dan Sigmoid. Kemudian diuji performanya menggunakan confusion matrix.

a. Senario Pembagian data 70:30

Pada percobaan skenario pertama didapatkan hasil tertinggi pada kernel Polinomial dengan akurasi sebesar 91.32%, menggunakan parameter  $C : 1, \gamma : 0.1, r : 2, d : 5$ , dengan 2 kelas

sentimen dan teknik smote. Tabel 5 dan Tabel 6 menunjukkan hasil dari pengujian skenario pertama.

Tabel 5. Hasil pengujian skenario pertama 3 kelas sentimen

Kernel	3 Kelas Sentimen		3 Kelas Sentimen	
	Tanpa SMOTE		Menggunakan SMOTE	
	Akurasi	Parameter	Akurasi	Parameter
Linear	73.07%	$C : 10$	85.36%	$C : 100$
RBF	73.77%	$C : 10, \gamma : 0.1$	89.41%	$C : 10, \gamma : 1$
Sigmoid	72.72%	$C : 100, \gamma : 0.01, r : 0$	85.58%	$C : 100, \gamma : 0.01, r : 0$
Polinomial	74.12%	$C : 1, \gamma : 0.001, r : 2, d : 10$	86.03%	$C : 1, \gamma : 0.001, r : 3, d : 8$

Tabel 6. Hasil pengujian skenario pertama 2 kelas sentimen

Kernel	2 Kelas Sentimen		2 Kelas Sentimen	
	Tanpa SMOTE		Menggunakan SMOTE	
	Akurasi	parameter	Akurasi	Parameter
Linear	84.61%	$C : 10$	91.05%	$C : 10$
RBF	86.36%	$C : 10, \gamma : 0.1$	91.05%	$C : 10, \gamma : 1$
Sigmoid	84.26%	$C : 100, \gamma : 0.1, r : 0$	90.51%	$C : 100, \gamma : 0.1, r : 2$
Polinomial	85.31%	$C : 0.1, \gamma : 0.1, r : 2, d : 10$	91.32%	$C : 1, \gamma : 0.1, r : 2, d : 5$

b. Skenario Pembagian data 80:20

Pada percobaan skenario kedua didapatkan hasil tertinggi pada kernel RBF dengan akurasi sebesar 91.86% , menggunakan parameter  $C : 0.1$

,  $\gamma : 1, r : 2, d : 5$  dengan 2 kelas sentimen dan teknik smote. Tabel 7 dan Tabel 8 menunjukkan hasil dari pengujian skenario kedua.

Tabel 7. Hasil pengujian skenario kedua 3 kelas sentimen

Kernel	3 Kelas Sentimen		3 Kelas Sentimen	
	Tanpa SMOTE		Menggunakan SMOTE	
	Akurasi	Parameter	Akurasi	Parameter
Linear	75.39%	$C : 10$	85.47%	$C : 100$
RBF	76.43%	$C : 10, \gamma : 0.1$	88.85%	$C : 100, \gamma : 0.1$
Sigmoid	75.39%	$C : 100, \gamma : 0.1, r : 0$	86.82%	$C : 100, \gamma : 0.1, r : 0$
Polinomial	75.91%	$C : 0.001, \gamma : 0.1, r : 3, d : 10$	86.14%	$C : 10, \gamma : 0.001, r : 2, d : 10$

Tabel 8. Hasil pengujian skenario kedua 2 kelas sentimen

Kernel	2 Kelas Sentimen		2 Kelas Sentimen	
	Tanpa SMOTE		Menggunakan SMOTE	
	Akurasi	Parameter	Akurasi	Parameter
Linear	84.29%	$C : 10$	91.46%	$C : 10$
RBF	84.81%	$C : 10, \gamma : 0.1$	90.24%	$C : 10, \gamma : 0.1$
Sigmoid	84.29%	$C : 100, \gamma : 0.1, r : 0$	91.46%	$C : 100, \gamma : 1, r : 0$
Polinomial	84.29%	$C : 1, \gamma : 0.01, r : 2, d : 5$	91.86%	$C : 0.1, \gamma : 1, r : 2, d : 5$

c. Skenario Pembagian data 90:10

Pada percobaan skenario ketiga didapatkan hasil tertinggi pada kernel Linear dengan akurasi sebesar 92.68% , menggunakan parameter  $C : 10$

dengan 2 kelas sentimen dan teknik smote. Tabel 9 dan Tabel 10 menunjukkan hasil dari pengujian skenario ketiga.



Tabel 9. Hasil pengujian skenario ketiga 3 kelas sentimen

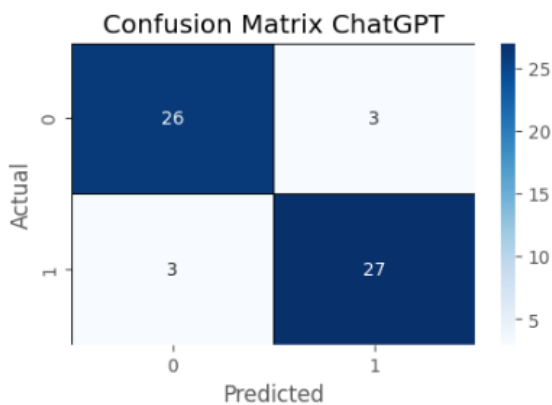
Kernel	2 Kelas Sentimen			
	Tanpa SMOTE		Menggunakan SMOTE	
	Akurasi	Parameter	Akurasi	Parameter
Linear	72.91%	C : 10	86.48%	C : 100
RBF	77.08%	C : 10, $\gamma$ : 0.1	91.89%	C : 100, $\gamma$ : 0.1
Sigmoid	79.12%	C : 100, $\gamma$ : 0.1, $r$ : 0	89.18%	C : 100, $\gamma$ : 0.1, $r$ : 0.1
Polinomial	78.16%	C : 0.001, $\gamma$ : 0.1, $r$ : 3, $d$ : 10	86.48%	C : 100, $\gamma$ : 0.1, $r$ : 2, $d$ : 10

Tabel 10. Hasil pengujian skenario ketiga 2 kelas sentimen

Kernel	2 Kelas Sentimen			
	Tanpa SMOTE		Menggunakan SMOTE	
	Akurasi	Parameter	Akurasi	Parameter
Linear	81.25%	C : 1	<b>92.68%</b>	<b>C : 10</b>
RBF	82.29%	C : 10, $\gamma$ : 0.01	91.86%	C : 100, $\gamma$ : 0.1
Sigmoid	82.33%	C : 100, $\gamma$ : 0.1, $r$ : 1	91.86%	C : 100, $\gamma$ : 0.11, $r$ : 0
Polinomial	83.29%	C : 0.001, $\gamma$ : 0.01, $r$ : 3, $d$ : 8	91.06%	C : 0.1, $\gamma$ : 1, $r$ : 1, $d$ : 8

Dari tiga skenario pengujian yang telah dilakukan, diperoleh bahwa model terbaik yaitu menggunakan kernel Linear dengan parameter C: 10 pada 2 kelas sentimen dan menggunakan teknik smote. Hasil evaluasi dari model tersebut memiliki nilai accuracy yang paling tinggi dari seluruh hasil pengujian yaitu sebesar 92.68%.

Selanjutnya hasil pemodelan terbaik tersebut diterapkan pada dataset baru dari ulasan aplikasi ChatGPT, Perplexity, dan Microsoft Bing Chat masing masing sebanyak 500 ulasan untuk melihat sentiment pengguna pada tiap aplikasinya.



Gambar 2. Hasil Cofusion Matrix ChatGPT

Hasil yang didapatkan dari pengujian ini mendapatkan nilai *accuracy* sebesar 89.83%. Selanjutnya dilakukan *Classification Report* dan

diperoleh nilai *presisi*, *recall*, dan *f1-score* dapat dilihat pada Tabel 11.

Tabel 11. Hasil Classification Report ChatGPT

Classification Report ChatGPT				
	Precision	Recall	F1-Score	Support
Negatif	0.90	0.90	0.90	<b>29</b>
Positif	0.90	0.90	0.90	<b>30</b>

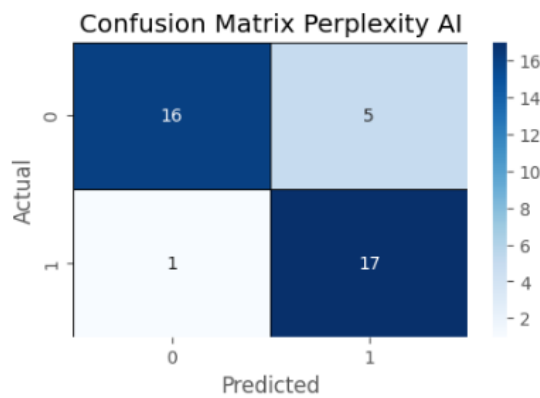
Tabel 11 menampilkan hasil penerapan best model pada dataset ChatGPT. Hasil evaluasi menunjukkan bahwa model SVM yang digunakan memiliki kinerja yang baik dalam mengklasifikasikan data. Akurasi, presisi, recall, dan F1-score semua mencapai nilai yang tinggi (sekitar 90%), menunjukkan bahwa model ini efektif dalam mengidentifikasi baik kelas positif maupun kelas negatif.



Gambar 3. Visualisasi Wordcloud ulasan pengguna ChatGPT

Berdasarkan Gambar 3 dapat dilihat bahwa ulasan ChatGPT didominasi pujian kepada

aplikasi karena bagus, mudah dan bermanfaat serta memberikan informasi yang dapat membantu penggunaanya dalam menyelesaikan tugas



Gambar 4. Hasil Cofusion Matrix Perplexity.AI

Hasil yang didapatkan dari pengujian ini mendapatkan nilai *accuracy* sebesar 84.61%. Selanjutnya dilakukan *Classification Report* dan diperoleh nilai *presisi*, *recall*, dan *f1-score* dapat dilihat pada tabel 12.

Tabel 12. Hasil Classification Report Perplexity.AI

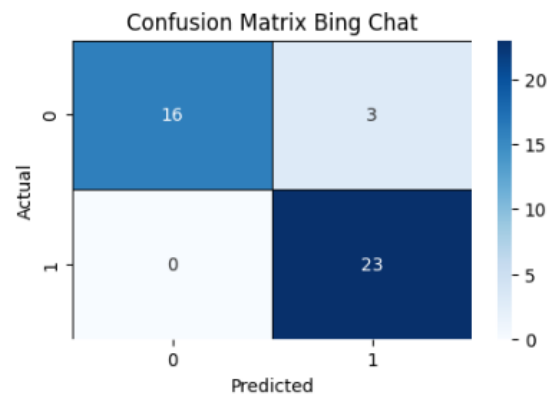
Classification Report Perplexity.AI				
	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
Negatif	0.94	0.76	0.84	<b>21</b>
Positif	0.77	0.94	0.85	<b>18</b>

Tabel 12 menampilkan hasil penerapan best model pada dataset Perplexity.Ai, didapatkan nilai akurasi sebesar 84.61%. Hasil evaluasi menunjukkan bahwa model SVM memiliki kinerja yang baik dalam mengklasifikasikan data, meskipun dengan beberapa perbedaan yang cukup signifikan antara presisi dan recall untuk kelas 0 dan kelas 1. Presisi untuk kelas 0 lebih tinggi (94%) dibandingkan dengan presisi untuk kelas 1 (77%), sedangkan recall untuk kelas 1 lebih tinggi (94%) dibandingkan dengan recall untuk kelas 0 (76%). F1-score untuk kedua kelas menunjukkan kinerja yang seimbang, dengan nilai yang cukup tinggi (di atas 0.8).



Gambar 5. Visualisasi Wordcloud ulasan pengguna Perplexity.AI

Berdasarkan Gambar 5 dapat dilihat bahwa ulasan Perplexity juga didominasi pujian kepada aplikasi karena bagus, dan sangat membantu penggunaanya.



Gambar 6. Hasil Cofusion Matrix Bing Chat

Hasil yang didapatkan dari pengujian ini mendapatkan nilai *accuracy* sebesar 92.85%. Selanjutnya dilakukan *Classification Report* dan diperoleh nilai *presisi*, *recall*, dan *f1-score* dapat dilihat pada tabel 13.

Tabel 13. Hasil Classification Report Bing Chat

Classification Report Bing Chat				
	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
Negatif	1.00	0.84	0.91	<b>19</b>
Positif	0.88	1.00	0.94	<b>23</b>

Tabel 13 menampilkan hasil penerapan best model pada dataset Microsoft Bing Chat. Hasil evaluasi menunjukkan bahwa model SVM memiliki kinerja yang sangat baik dalam mengklasifikasikan data. Akurasi, presisi, recall, dan F1-score semua mencapai nilai yang tinggi (di atas 0.9), menunjukkan bahwa model ini efektif dalam mengidentifikasi baik kelas positif maupun kelas negatif. Presisi untuk kelas negatif sangat



tinggi (100%), sedangkan untuk kelas positif cukup tinggi (88%). Recall untuk kelas positif juga sangat tinggi (100%), sedangkan untuk kelas negatif sedikit lebih rendah (84%)



Gambar 7. Visualisasi Wordcloud ulasan pengguna Microsoft Bing Chat

Dapat dilihat pada Gambar 6 bahwa ulasan pengguna Microsoft Bing Chat juga tidak berbeda jauh dari dua aplikasi sebelumnya. Kalimat pujian yang mengatakan bahwasanya bing merupakan aplikasi yang bagus dan membantu mendominasi pada wordcloud.

## 5 Kesimpulan

Berdasarkan hasil percobaan yang dilakukan untuk membandingkan evaluasi kernel SVM dalam analisis sentimen *chatbot* AI pada ulasan Google Play Store, kernel linier mendapat nilai akurasi tertinggi sebesar 92.68% pada splitting data 90:10 dengan 2 kelas sentimen serta menggunakan teknik SMOTE oversampling. Dari hasil penelitian tersebut terbukti bahwa kernel pada metode SVM dapat meningkatkan nilai akurasi dan penggunaan SMOTE oversampling berhasil menaikkan kinerja algoritma. Berdasarkan hasil pelabelan sentimen yang dilakukan pada ketiga *chatbot* AI menghasilkan jumlah sentimen positif lebih besar daripada sentimen negatif. Hal ini membuktikan bahwa pengguna sangat senang dan terbantu dengan adanya aplikasi *chatbot* AI.

## References

- Ananda, D., & Suryono, R. R. (2024). Analisis Sentimen Publik Terhadap Pengungsi Rohingya di Indonesia dengan Metode Support Vector Machine dan Na<sup>A</sup> ve Bayes. *Jurnal Media Informatika Budidarma*, 8(2), 748–757.
- Auliaddina, S., & Arifin, T. (2024). Use of Augmentation Data and Hyperparameter Tuning in Batik Type Classification using the CNN

- Model. *Sistemasi: Jurnal Sistem Informasi*, 13(1), 114–128.
- Caramiaux, B., & Fdili Alaoui, S. (2022). “Explorers of Unknown Planets”: Practices and Politics of Artificial Intelligence in Visual Arts. *Proc. ACM Hum.-Comput. Interact.*, 6(CSCW2). <https://doi.org/10.1145/3555578>
- Ernawati, S., & Wati, R. (2024). Evaluasi Performa Kernel SVM dalam Analisis Sentimen Review Aplikasi ChatGPT Menggunakan Hyperparameter dan VADER Lexicon. *Jurnal Buana Informatika*, 15(1), 40–49.
- Gezici, B., Bölücü, N., Tarhan, A., & Can, B. (2019). Neural Sentiment Analysis of User Reviews to Predict User Ratings. *2019 4th International Conference on Computer Science and Engineering (UBMK)*, 629–634. <https://doi.org/10.1109/UBMK.2019.8907234>
- Iorliam, A., & Ingio, J. A. (2024). A Comparative Analysis of Generative Artificial Intelligence Tools for Natural Language Processing. *Journal of Computing Theories and Applications*, 1(3), 311–325. <https://doi.org/10.62411/jcta.9447>
- Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61(4), 577–586. <https://doi.org/https://doi.org/10.1016/j.bushor.2018.03.007>
- Jiang, Y., Li, X., Luo, H., Yin, S., & Kaynak, O. (2022). Quo vadis artificial intelligence? *Discover Artificial Intelligence*, 2(1), 4. <https://doi.org/10.1007/s44163-022-00022-8>
- Kavak, A. C. (2023). ChatGPT, Google Bard, Microsoft Bing, Claude, and Perplexity: Which is the Right AI Tool? In *Zeo.org*. <https://zeo.org/resources/blog/chatgpt-google-bard-microsoft-bing-claude-and-perplexity-which-is-the-right-ai-to>
- Kumar, P., & Prusty, M. (2020). Outlier-SMOTE: A refined oversampling technique for improved detection of COVID-19. *Intelligence-Based Medicine*, 3–4, 100023. <https://doi.org/10.1016/j.ibmed.2020.100023>
- Liu, B. (2022). *Sentiment analysis and opinion mining*. Springer Nature.
- Nanda, M. A., Seminar, K. B., Nandika, D., & Maddu, A. (2018). A comparison study of kernel functions in the support vector machine and its application for termite detection. *Information*, 9(1), 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 SE-Information

- Systems Engineering Articles).  
<https://doi.org/10.29207/resti.v6i2.3906>
- Permata Aulia, T. M., Arifin, N., & Mayasari, R. (2021). Perbandingan Kernel Support Vector Machine (Svm) Dalam Penerapan Analisis Sentimen Vaksinasi Covid-19. *SINTECH (Science and Information Technology) Journal*, 4(2), 139–145. <https://doi.org/10.31598/sintechjournal.v4i2.762>
- Prastyo, P. H., Ardiyanto, I., & Hidayat, R. (2020a). Indonesian Sentiment Analysis: An Experimental Study of Four Kernel Functions on SVM Algorithm with TF-IDF. *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy, ICDABI 2020*. <https://doi.org/10.1109/ICDABI51230.2020.9325685>
- Prastyo, P. H., Ardiyanto, I., & Hidayat, R. (2020b). Indonesian Sentiment Analysis: An Experimental Study of Four Kernel Functions on SVM Algorithm with TF-IDF. *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*, 1–6. <https://doi.org/10.1109/ICDABI51230.2020.9325685>
- Rabbani, S., Safitri, D., Rahmadhani, N., & Anam, M. K. (2023). Perbandingan Evaluasi Kernel SVM untuk Klasifikasi Sentimen dalam Analisis Kenaikan Harga BBM: Comparative Evaluation of SVM Kernels for Sentiment Classification in Fuel Price Increase Analysis. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 3(2), 153–160.
- Rianti, D. L., Umaidah, Y., & Voutama, A. (2021). Tren Marketplace Berdasarkan Klasifikasi Ulasan Pelanggan Menggunakan Perbandingan Kernel Support Vector Machine. *STRING (Satuan Tulisan Riset Dan Inovasi Teknologi)*, 6(1), 98–105. <https://doi.org/10.30998/string.v6i1.9993>
- Ribas, J. (2023). *Building the New Bing*. Bing Blogs. <https://blogs.bing.com/search-quality-insights/february-2023/Building-the-New-Bing/>
- Rudd, J. M. (2018). Application of support vector machine modeling and graph theory metrics for disease classification. *Model Assisted Statistics and Applications*, 13(4), 341–349.
- Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbot: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI rush and its impact on higher education. *Journal of Applied Learning & Teaching*, 6(1), 1–9. <https://doi.org/https://doi.org/10.37074/jalt.2023.6.1.23> Abstract
- Saeidnia, H. R., Hosseini, E., Abdoli, S., & Ausloos, M. (2024). Unleashing the power of AI: a systematic review of cutting-edge techniques in AI-enhanced scientometrics, webometrics and bibliometrics. *Library Hi Tech, ahead-of-p*(ahead-of-print). <https://doi.org/10.1108/LHT-10-2023-0514>
- Sari, P. K., & Suryono, R. R. (2024). Komparasi Algoritma Support Vector Machine dan Random Forest untuk Analisis Sentimen Metaverse. *Jurnal Mnemonic*, 7(1), 31–39.
- Sengkey, D. F., Jacobus, A., & Manoppo, F. J. (2020). Effects of kernels and the proportion of training data on the accuracy of svm sentiment analysis in lecturer evaluation. *IAES International Journal of Artificial Intelligence*, 9(4), 734–743. <https://doi.org/10.11591/ijai.v9.i4.pp734-743>
- Sullivan, M. (2023). *Is Perplexity AI showing us the future of search?* <https://www.fastcompany.com/90883562/is-perplexity-ai-showing-us-the-future-of-search>
- Tuttle, J. F., Blackburn, L. D., & Powell, K. M. (2020). On-line classification of coal combustion quality using nonlinear SVM for improved neural network NOx emission rate prediction. *Computers & Chemical Engineering*, 141, 106990.
- Wahyu Handani, S., Intan Surya Saputra, D., Hasirun, Mega Arino, R., & Fiza Asyrofi Ramadhan, G. (2019). Sentiment analysis for go-jek on google play store. *Journal of Physics: Conference Series*, 1196(1). <https://doi.org/10.1088/1742-6596/1196/1/012032>
- Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q.-L., & Tang, Y. (2023). A Brief Overview of ChatGPT: The History, Status Quo and Potential Future Development. *IEEE/CAA Journal of Automatica Sinica*, 10(5), 1122–1136. <https://doi.org/10.1109/JAS.2023.123618>
- Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., Qiu, C.-W., Qiu, J., Hua, K., Su, W., Wu, J., Xu, H., Han, Y., Fu, C., Yin, Z., Liu, M., ... Zhang, J. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation*, 2(4). <https://doi.org/10.1016/j.xinn.2021.100179>
- Zirar, A., Ali, S. I., & Islam, N. (2023). Worker and workplace Artificial Intelligence (AI) coexistence: Emerging themes and research agenda. *Technovation*, 124, 102747. <https://doi.org/https://doi.org/10.1016/j.technova.2023.102747>

