

Evaluasi Performa Kernel SVM dalam Analisis Sentimen Review Aplikasi ChatGPT Menggunakan Hyperparameter dan VADER Lexicon

Siti Ernawati^{*1}, Risa Wati²

¹Program Studi Sistem Informasi, Fakultas Teknologi Informasi, Universitas Nusa Mandiri, Jakarta Timur 13620, Indonesia

²Program Studi Sistem Informasi, Universitas Bina Sarana Informatika, Jakarta Pusat 10450, Indonesia
Email: siti.ste@nusamandiri.ac.id, risawati.rwx@bsi.ac.id

Abstract. *Kernel SVM Performance Evaluation on Sentiment Analysis of ChatGPT Application Review using Hyperparameter and VADER Lexicon.* ChatGPT is an artificial intelligence language model that responds to user questions and statements. ChatGPT has benefits and drawbacks for users. This raises comments on social media about whether ChatGPT is useful. This research is about sentiment analysis of ChatGPT application reviews using linear kernel SVM, RBF, polynomial, and sigmoid. Labeling uses VADER lexicon and hyperparameter to produce the best parameters. The research objectives are to determine whether the ChatGPT application can provide benefits and prove whether the kernel in SVM can increase the accuracy value. The percentage distribution between test data and training data is 70:30, 80:30, and 90:10. Preprocessing is done and then grouped into positive and negative reviews. A hyperparameter is performed on the C and Gamma parameters to produce the maximum value. Experimental results obtained the highest accuracy using SVM kernel RBF scenario 90:10 with an accuracy value of 92.72%, precision 92.44%, f1-score 96.10%, and AUC 88%.

Keywords: SVM, kernel, hyperparameter, VADER lexicon, sentiment analysis

Abstrak. *Evaluasi Performa Kernel SVM dalam Analisis Sentimen Review Aplikasi ChatGPT Menggunakan Hyperparameter dan VADER Lexicon.* ChatGPT merupakan model bahasa kecerdasan buatan yang merespon pertanyaan dan pernyataan pengguna. ChatGPT memiliki manfaat dan kelemahan bagi pengguna. Hal ini menimbulkan komentar pada media sosial tentang manfaat dari ChatGPT. Penelitian ini membahas tentang analisis sentimen review aplikasi ChatGPT menggunakan SVM kernel linier, RBF, polinomial dan sigmoid. Pelabelan menggunakan VADER lexicon dan hyperparameter untuk menghasilkan parameter terbaik. Tujuan penelitian yaitu apakah aplikasi ChatGPT dapat memberikan manfaat dan membuktikan apakah kernel pada SVM dapat meningkatkan nilai akurasi. Diskenariokan persentase pembagian antara data uji dan data latih adalah 70:30, 80:30, dan 90:10. Setelah dilakukan preprocessing, kemudian dikelompokkan menjadi review positif dan negatif. Dilakukan hyperparameter terhadap parameter C dan Gamma sehingga menghasilkan nilai maksimal. Hasil eksperimen diperoleh akurasi tertinggi menggunakan SVM kernel RBF skenario 90:10 dengan nilai accuracy 92.72%, precision 92.44%, f1-score 96.10% dan AUC 88%.

Kata Kunci: SVM, kernel, hyperparameter, VADER lexicon, analisis sentimen

1. Pendahuluan

ChatGPT (*Generative Pre-trained Transformer*) adalah model bahasa kecerdasan buatan (AI) yang mulai diperkenalkan pada November 2022. ChatGPT merupakan salah satu iterasi dari keluarga model bahasa GPT yang dikembangkan oleh OpenAI. Ini adalah varian dari model GPT yang dimaksudkan untuk berinteraksi dengan orang lain dalam bentuk chat atau percakapan. ChatGPT memberikan respons percakapan yang dihasilkan untuk permintaan pertanyaan [1]. ChatGPT merespon pertanyaan dan pernyataan dari pengguna sesuai dengan konteks dalam berbagai jenis percakapan. Menghasilkan respon yang lancar dan berkualitas tinggi terhadap pertanyaan manusia [2]. Peluncuran ChatGPT telah mengungkap berbagai kemungkinan bahwa *Large Language Models* (LLMs) dapat menggantikan kecerdasan manusia [3]. ChatGPT

merupakan teknologi AI yang memiliki banyak manfaat penggunaan di media sosial seperti otomatisasi layanan, pembuatan konten, chatbot, analisis sentimen, dan pengumpulan data. Tetapi memiliki beberapa kelemahan seperti kurangnya empati, potensi bias dan ketidakakuratan, kesadaran konteks yang terbatas, kesulitan dalam penskalaan, dan kebutuhan pengawasan manusia [4]. Hal ini menimbulkan banyak *review* pada media sosial Twitter terkait apakah ChatGPT membantu menjawab pertanyaan yang diajukan oleh pengguna aplikasi ChatGPT.

Dalam era digital dan informasi yang semakin berkembang, analisis sentimen menjadi penting karena banyaknya data teks yang dihasilkan oleh pengguna internet, seperti ulasan produk pada platform *e-commerce*, komentar di media sosial, artikel berita, dan lain-lain. Analisis sentimen ini membantu organisasi, bisnis, bahkan peneliti untuk memahami bagaimana masyarakat merasakan dan merespons berbagai hal yang sedang terjadi. Banyaknya opini yang beredar dapat menciptakan kekacauan informasi, sehingga sulit membedakan antara informasi yang benar dan salah, serta dapat mempengaruhi persepsi masyarakat. Oleh karena itu diadakan penelitian dengan objek dari penelitian yaitu kumpulan *review* yang diutarakan masyarakat melalui media sosial. *Review* ini terus meningkat seiring berjalannya waktu dan dapat dimanfaatkan oleh peneliti untuk mengetahui sentimen masyarakat dengan cara melakukan analisis terhadap *review* tersebut.

Untuk melakukan analisis data pada analisis regresi dan klasifikasi, algoritma SVM (*Support Vector Machine*) digunakan sebagai model pengajaran supervisi. Dalam praktiknya, algoritma SVM banyak digunakan sebagai pengklasifikasi [5]. Hasil akhir klasifikasi pada algoritma SVM dipengaruhi oleh pemilihan kernel [6]. Fungsi kernel pada algoritma SVM memainkan peran penting dalam meningkatkan kinerja algoritma SVM [7]. Terdapat empat kernel yang sering digunakan pada algoritma SVM yaitu linier, sigmoid, polinomial dan RBF. Karena algoritma SVM memberikan hasil klasifikasi yang cukup akurat, maka algoritma SVM dengan kernel RBF sering digunakan [8].

Objek dari penelitian yaitu kumpulan *review* yang diutarakan masyarakat melalui media sosial. *Review* ini terus meningkat seiring berjalannya waktu dan dapat dimanfaatkan oleh peneliti untuk mengetahui sentimen masyarakat dengan cara melakukan analisis terhadap *review* tersebut. Tujuan dari penelitian ini yaitu untuk mengetahui manfaat ChatGPT bagi penggunaannya, serta membuktikan bahwa kernel pada SVM dapat meningkatkan nilai akurasi pada *review* Aplikasi ChatGPT.

Pada penelitian ini akan melakukan analisis sentimen *review* aplikasi ChatGPT menggunakan algoritma SVM. VADER Lexicon digunakan untuk mengklasifikasikan *review* kedalam *review* positif dan *review* negatif. VADER Lexicon, juga dikenal sebagai Dictionary Aware and Sentiment Reasoner, adalah metode analisis sentimen berbasis kamus leksikon yang terbukti berhasil dalam memeriksa teks berbahasa alami [9]. VADER Lexicon ini berfungsi untuk mengevaluasi sentimen dalam bentuk teks seperti *review* atau dokumen dan menentukan sentimen tersebut termasuk dalam kategori positif, negatif, atau netral. Dalam penelitian ini menggunakan hyperparameter untuk meningkatkan akurasi. Hyperparameter digunakan untuk mengoptimalkan parameter agar mendapatkan kinerja terbaik [10].

2. Tinjauan Pustaka

2.1. Analisis Sentimen

Analisis sentimen merupakan proses menganalisis teks untuk menentukan sentimen atau perasaan yang terkandung dalamnya. Semakin populernya internet telah membuat web menjadi sumber utama penghasil informasi. Banyak pengguna menggunakan berbagai sumber daya *online* untuk mengekspresikan pandangan dan pendapat [14]. Prosesnya mencakup beberapa operasi yang berakhir dengan klasifikasi apakah teks tertentu mengungkapkan sentimen positif atau negatif.

2.2. ChatGPT

ChatGPT diperkenalkan sebagai prototipe pada bulan November 2022 telah menarik minat para *engineer*, pengguna media sosial, pemilik bisnis, penulis, dan mahasiswa [15]. *Chatbot*

ini dapat berkomunikasi seperti manusia karena dikembangkan menggunakan Open AI. ChatGPT dirancang untuk menghasilkan teks seperti manusia, memungkinkan untuk terlibat dalam percakapan dengan pengguna secara alami [16]. ChatGPT tidak hanya dapat memahami bahasa Inggris, tetapi juga bahasa lainnya dalam tingkat yang bervariasi. Selain itu, ChatGPT hanya mengolah informasi yang diberikan atau diinputkan dan menghasilkan tanggapan berdasarkan pola yang telah pelajari.

2.3. Support Vector Machine (SVM)

Support Vector Machine (SVM) adalah teknik penting dan merupakan teknik dasar dalam *machine learning* [17]. SVM adalah salah satu pembelajaran mesin yang banyak digunakan yaitu teknik untuk mendeteksi polaritas dari sebuah teks [18]. Algoritma ini sangat populer karena memiliki akurasi, presisi, dan kualitas yang tinggi [19]. SVM memiliki kemampuan untuk memprediksi regresi dan klasifikasi. Dalam mengimplementasikan kinerja model, penelitian ini menggunakan SVC (*Support Vector Classification*) *library sklearn* yaitu implementasi *libsvm* [20] pada Python.

Cara kerja SVM yaitu mencari vektor pendukung untuk setiap kelas. Sampel kelas yang paling dekat dengan sampel kelas lainnya disebut *support vector*. SVM menghitung margin setelah mendapatkan *support vector*. Dengan menggunakan *support vector*, margin dapat dianggap sebagai jalan yang memisahkan dua kelas. *Support vector* berfungsi sebagai batas tepi jalan, juga disebut bahu jalan dan SVM mencari margin terbesar atau jalan terlebar yang mampu memisahkan dua kelas.

2.4. Valence Aware Dictionary and Sentiment Reasoner (VADER Lexicon)

VADER merupakan alat analisis sentimen yang khusus digunakan dalam pemrosesan bahasa alami untuk mengukur sentimen atau perasaan dalam teks. Alat ini dikembangkan oleh tim peneliti di SocialCog dan dirilis sebagai bagian dari pustaka Python yang disebut NLTK (*Natural Language Toolkit*). VADER adalah alat analisis sentimen berbasis leksikon dan aturan yang secara khusus disesuaikan dengan sentimen yang diungkapkan di media sosial [21]. VADER akan menghasilkan skor yang dinyatakan dengan *compound* yaitu gabungan dari masing-masing skor positif, negatif, dan netral. Nilai di atas nol cenderung menunjukkan sentimen positif, nilai di bawah nol cenderung menunjukkan sentimen negatif, dan nilai mendekati nol mengindikasikan sentimen netral. Setiap kata akan dihitung untuk mengetahui nilai negatif dan positif [19]. Label kelas sentimen secara otomatis diproses ke VADER Lexicon pada Python. *Library* NLTK menggunakan modul VADER Lexicon pada Python, yang menyediakan kamus kumpulan kata dengan nilai yang dapat dikategorikan. Modul ini digunakan untuk melabelkan data pada setiap sentimen. Setiap kata yang ditemukan dalam data akan termasuk dalam salah satu dari tiga kategori, yaitu positif, negatif, dan netral.

2.5. GridsearchCV

Salah satu teknik untuk mengoptimalkan nilai akurasi adalah dengan menggunakan *grid search* [22]. *Grid search* mempunyai tujuan untuk menentukan hasil kombinasi dari parameter yang terbaik untuk kinerja model yang diusulkan [23]. GridSearchCV merupakan sebuah fungsi atau objek dalam pustaka (*library*) Python yang disebut scikit-learn (*sklearn*). *Library* scikit-learn merupakan salah satu pustaka populer untuk *machine learning* dan *data mining* dalam Python. GridSearchCV digunakan untuk melakukan pencarian parameter terbaik dari model *machine learning* dengan cara yang sistematis.

3. Metodologi Penelitian

Gambar 1 menunjukkan beberapa langkah yang harus dilakukan untuk melakukan eksperimen penelitian ini.



Gambar 1. Alur Penelitian

3.1. Dataset

Data yang digunakan adalah *review* terhadap aplikasi chatGPT yang diambil dari <https://www.kaggle.com/datasets/furqanrustam118/chatgpt-tweets> [24]. *Review* ChatGPT dipilih karena fenomena platform ChatGPT saat ini yang sedang menjadi perhatian dunia teknologi [25]. ChatGPT memiliki banyak fungsi seperti menerjemahkan bahasa, membantu *programmer* menyelesaikan masalah *source code*, menjelaskan ulang konsep dengan bentuk sederhana, membuat draf sebuah artikel, membantu *Customer Service* melayani pelanggan, menyediakan informasi dan lain-lain [26]. Data yang ada akan diubah menjadi bentuk *dataframe* dengan menggunakan *library* *pandas*.

3.2. Preprocessing

Data yang telah diubah menjadi bentuk *dataframe* akan dilakukan *preprocessing*. Tahapan *preprocessing* yang dilakukan yaitu *case folding*, *cleaning text*, *tokenization*, *lemmatization*. Tujuan dari langkah *preprocessing* tidak hanya untuk mendapatkan analisis yang lebih baik tetapi untuk mengurangi dimensi *input* data karena banyak kata yang tidak berguna dan harus dihilangkan karena tidak berdampak pada polaritas teks.

3.3. Labeling with VADER Lexicon

Pelabelan yang dilakukan menggunakan VADER Lexicon. VADER merupakan *library* analisis sentimen yang sangat populer dalam pemrosesan bahasa alami yang dirancang khusus untuk mendeteksi sentimen positif, negatif, atau netral dalam teks. Namun dalam penelitian ini hanya mendeteksi dua sentiment, yaitu positif dan negatif. Hal ini dilakukan karena untuk memaksimalkan evaluasi yang nantinya akan dilakukan pada akhir tahapan penelitian.

3.4. TF-IDF

Proses untuk mengekstrak fitur (*feature extraction*), pada tahapan ini akan terlihat seberapa pentingnya kata-kata yang muncul dalam sebuah dokumen.

3.5. Hyperparameter with GridsearchCV

Pada tahap ini, proses dilakukan untuk menemukan kombinasi *hyperparameter* yang akan menghasilkan model yang berfungsi dengan baik menggunakan teknik GridsearchCV menggunakan *library* *scikit-learn* yang sangat berguna untuk melakukan pencarian grid (*grid search*) terhadap berbagai kombinasi *hyperparameter* yang berbeda untuk model pembelajaran mesin. Hyperparameter dilakukan terhadap parameter C dan Gamma pada model klasifikasi yang diusulkan. Tujuannya adalah untuk mengidentifikasi kombinasi *hyperparameter* yang memberikan kinerja terbaik bagi model yang diusulkan.

3.6. Classification with SVM

Tahapan ini dilakukan untuk membuat model klasifikasi dengan persentase pembagian data latih dan data uji akan diskenariokan menjadi tiga dengan rincian yaitu 70:30, 80:30 dan 90:10 dengan membandingkan empat kernel yang ada diantaranya yaitu RBF, linier, polinomial dan sigmoid.

3.7. Evaluation

Tahapan terakhir yaitu melakukan perhitungan *confusion matrix* untuk menghasilkan nilai *accuracy*, *precision*, *recall* dan F1-Score. Selain menggunakan *confussion matrix*, juga menggunakan kurva ROC (*Receiver Operator Characteristic*) dan nilai AUC (*Area Under*

Curve). Recall, precision, dan F-measure merupakan cara umum dalam mengevaluasi hasil eksperimen machine learning. Ukuran-ukurannya diberi nama berdasarkan asal usul dalam pengambilan informasi dan menyajikan bias tertentu. ROC telah dijadikan standar evaluasi dan penetapan standar untuk membandingkan true positive rate dan false positive rate [27]. Pada penelitian ini yang menjadi metrik utama untuk analisis performa kernel yaitu accuracy karena accuracy, sensitivity, specificity, positive predictive value dan F1 score diukur pada satu ambang batas yang optimal [25].

4. Hasil dan Diskusi

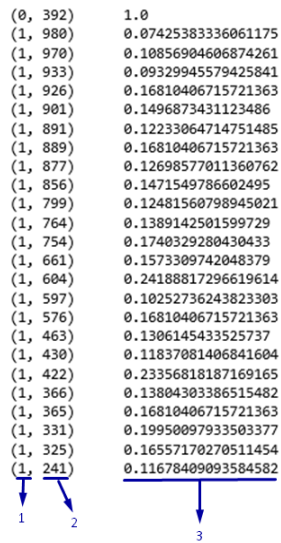
4.1. Preprocessing dan Feature Extraction

Preprocessing dilakukan untuk membersihkan dan mengubah teks mentah menjadi format yang lebih mudah untuk diproses oleh algoritma pembelajaran mesin dan mudah untuk dianalisa. Tabel 1 menunjukkan hasil preprocessing dalam penelitian yang dilakukan, dengan menggunakan beberapa tahapan yaitu case folding, cleaning text, tokenization dan lemmatization.

Tabel 1. Hasil Preprocessing dalam Penelitian

Data	Case Folding	Cleaning Text	Tokenization	Lemmatization
I installed this on my iPhone and it works great. I have already been using chat GPT through Google. So I had to use my Google log on on this iPhone app. I clicked on my Google account, and it filled in the password using Authenticator. But the password was wrong! so I typed in the password. Then it authenticated me on Google on my iPad. I confirmed that I was trying to login on my iPhone and all is good.	i installed this on my iphone and it works great. i have already been using chat gpt through google. so i had to use my google log on on this iphone app. i clicked on my google account, and it filled in the password using authenticator. but the password was wrong! so i typed in the password. then it authenticated me on google on my ipad. i confirmed that i was trying to login on my iphone and all is good.	installed this iphone works great have already been using chat through google google this iphone clicked google account filled password using authenticator password wrong typed password then authenticated google ipad confirmed that trying login iphone good	['installed', 'this', 'iphone', 'works', 'great', 'have', 'already', 'been', 'using', 'chat', 'through', 'google', 'google', 'this', 'iphone', 'clicked', 'google', 'account', 'filled', 'password', 'using', 'authenticator', 'password', 'wrong', 'typed', 'password', 'then', 'authenticated', 'google', 'ipad', 'confirmed', 'that', 'trying', 'login', 'iphone', 'good']	installed this iphone work great have already been using chat through google google this iphone clicked google account filled password using authenticator password wrong typed password then authenticated google ipad confirmed that trying login iphone good

TF-IDF merupakan salah satu teknik feature extraction yang digunakan dalam penelitian. TF-IDF sangat berguna dalam mengidentifikasi kata-kata yang penting dalam sebuah dokumen dan dapat membantu dalam pemrosesan teks. Gambar 2 menunjukkan hasil pengolahan feature extraction menggunakan teknik TF-IDF menggunakan Python. Berdasarkan Gambar 2 disimpulkan bahwa kode 1 menunjukkan nomor baris dari setiap data yang diolah. Kode 2 merupakan nomor integer unik setiap kata yang ada pada baris. Kode 3 merupakan hasil pembobotan (skor) yang dihitung menggunakan teknik TF-IDF.



Gambar 2. Hasil Pengolahan Feature Extraction Menggunakan Teknik TF-IDF

4.2. Hasil Labelling with VADER Method

Hasil klasifikasi menggunakan VADER Lexicon diperoleh jumlah *review* positif sebanyak 1263, *review* negatif sebanyak 208 dan *review* netral sebanyak 587. Hasil dari pelabelan terhadap *review* penggunaan aplikasi ChatGPT menggunakan VADER Lexicon. Gambar tersebut menunjukkan persentase 61,37% *review* masuk ke dalam kategori positif, 10,11% masuk ke dalam kategori negatif dan 28,52% masuk ke dalam kategori netral.

Tabel 2 merupakan hasil perhitungan skor sentimen menggunakan VADER Lexicon. Pada Tabel 2 terlihat skor dari masing-masing sentiment, yaitu skor negatif, positif, netral, *compound*, dan hasil dari *polarity*.

Tabel 2. Hasil Pelabelan Menggunakan VADER Lexicon

Review	Compound	Negative	Positive	Neutral	Polarity
very nice official with fake one store chatgpt iphone ipad downside lack ipad optimized screen size sure plan added later even with that little change might become five five assuming they other feature that browser version like browsing plug they that will certainly place opposed browser that keep open constantly nice love vibration iphone speech text feature seems work better than default apple occasionally long wait long query great first look forward improvement	0,9716	0,076	0,31	0,614	1
installed this iphone work great have already been using chat through google google this iphone clicked google account filled password using authenticator password wrong typed password then authenticated google ipad confirmed that trying login iphone good tried install ipad again asked login unfortunately even though typed password iphone authenticatedid update password file type password again unfortunately ipad this doe fill full screen look like only meant iphone insult injury number appear keyboard they blank have click button view password guess where proper number finally password typed incorrectly clicked transfer error message whole lock guess will only chat iphone guy need figure this stuff	-0,6808	0,138	0,1	0,762	0

4.3. Proses *Hyperparameter*

Pada proses ini mencari kombinasi optimal dari nilai-nilai hyperparameter untuk meningkatkan kinerja model. Pada Gambar3, penggunaan GridSearchCV dari *library* scikit-learn dalam Python untuk mencari kombinasi hyperparameter terbaik untuk model *Support Vector Machine* (SVM) dengan kernel *Radial Basis Function* (RBF). Gambar 4 menampilkan parameter terbaik dari hasil hyperparameter menggunakan GridsearchCV, dengan hasil C=10, Gamma=0.1 dan Kernel=RBF.

```

from sklearn.model_selection import GridSearchCV
from sklearn.svm import SVC

# membangun model dengan parameter C, gamma, dan kernel
model = SVC()
parameters = {
    'kernel': ['rbf'],
    'C': [1, 10, 100, 1000],
    'gamma': [0.01, 0.1, 1, 10, 100]
}
grid_search = GridSearchCV(model, parameters)

# melatih model dengan fungsi fit
grid_search.fit(train_X_tfidf, train_Y)

GridSearchCV(estimator=SVC(),
              param_grid={'C': [1, 10, 100, 1000],
                          'gamma': [0.01, 0.1, 1, 10, 100], 'kernel': ['rbf']})
    
```

Gambar 3. Penggunaan GridSearchCV dalam Python

```

# menampilkan parameter terbaik dari objek grid_search
print(grid_search.best_params_)

{'C': 10, 'gamma': 0.1, 'kernel': 'rbf'}
    
```

Gambar 4. Menampilkan hasil Parameter Terbaik dari GridsearchCV

4.4. Hasil Klasifikasi SVM

Tabel 3 menunjukkan kinerja terbaik dari model SVM dengan kernel yang digunakan sesuai dengan skenario pembagian *data training* dan *data testing*.

1. Skenario Pembagian Data 70:30

Pada skenario pertama dapat dilihat bahwa kernel yang tertinggi yaitu RBF dengan nilai parameter C=10, Gamma = 0,1, *accuracy* = 92,07%, *precision* = 92,44%, *F1_score* = 95,74% dan untuk nilai AUC = 79%. Tabel 3 menunjukkan persamaan hasil skenario pembagian data 70:30.

Tabel 3. Hasil Eksperimen Skenario Pembagian Data 70:30

Kernel	C	Gamma	Accuracy	Precision	F1_Score	AUC
RBF	10	0,1	92,07%	92,44%	95,74%	79%
Linier	1	0,01	91,42%	91,27%	95,43%	79%
Polinomial	1	1	89,97%	89,94%	94,70%	72%
Sigmoid	10	0,1	91,42%	91,27%	95,43%	79%

2. Skenario Pembagian Data 80:20

Pada skenario kedua, terlihat banyak nilai yang sama terhadap hasil eksperimen yang dilakukan. Seperti *accuracy*, *precision*, *F1_Score* dan AUC. Tabel 4 menunjukkan persamaan hasil skenario pembagian data 80:20.

Tabel 4. Hasil Eksperimen Skenario Pembagian Data 80:20

Kernel	C	Gamma	Accuracy	Precision	F1_Score	AUC
RBF	10	0,1	91,26%	91,48%	95,30%	83%
Linier	1	0,01	91,26%	91,27%	95,31%	83%
Polinomial	1	1	89,56%	89,51%	94,47%	80%
Sigmoid	10	0,1	91,26%	91,27%	95,31%	83%

3. Skenario Pembagian Data 90:10

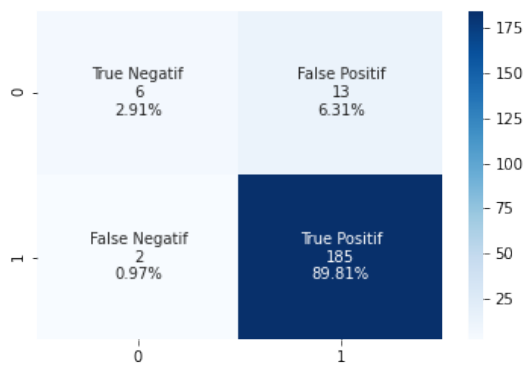
Pada skenario terakhir yaitu pembagian data 90:10 didapatkan hasil maksimal pada kernel RBF dengan $C = 10$, $\text{Gamma} = 0.1$, $\text{accuracy} = 92.72\%$, $\text{precision} = 93.43\%$, $F_Score = 96.10\%$, $\text{AUC} = 88\%$. Tabel 5 menunjukkan persamaan hasil skenario pembagian data 90:10.

Tabel 5. Hasil Eksperimen Skenario Pembagian Data 90:10

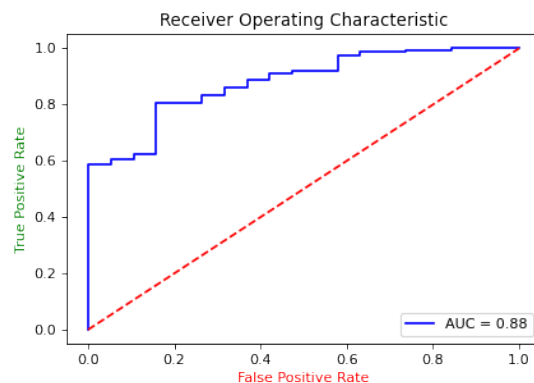
Kernel	C	Gamma	Accuracy	Precision	F1_Score	AUC
RBF	10	0,1	92,72%	93,43%	96,10%	88%
Linier	1	0,01	92,23%	92,12%	95,90%	86%
Polinomial	1	1	90,78%	90,78%	95,17%	80%
Sigmoid	1	1	92,23%	92,12%	95,90%	80%

4.5. Evaluasi

Dengan menggunakan *confusion matrix*, evaluasi kinerja model yang diusulkan dilakukan dengan melihat hasil *accuracy*, *precision*, dan *F1-Score*. *Confusion matrix* dihitung berdasarkan nilai *true positive* (tp), *false positive* (fp), *true negative* (tn), dan *false negative* (fn) sesuai kelas yang ditetapkan. Sedangkan *precision* digunakan untuk memprediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Evaluasi dilakukan terhadap nilai maksimal yang diperoleh dari tiga skenario, yaitu pada skenario ketiga dengan hasil *true positive* (tp) = 89,81%, *false positive* (fp) = 6,31%, *true negative* (tn) = 2,91%, dan *false negative* (fn) = 0,97%. Dengan nilai *accuracy* sebesar 92,72%. *Confusion matrix* dapat dilihat pada Gambar 5 dan nilai *accuracy* dapat dilihat pada Persamaan 1.

**Gambar 5. Confusion Matrix**

$$\text{Accuracy} = \frac{(89,81+2,91)}{(89,81+6,31+0,97+2,91)} = 92,72 \quad (1)$$



Gambar 6. Kurva ROC Hasil Eksperimen Skenario Ketiga

Gambar 6 menggambarkan kurva ROC yang menunjukkan *confusion matrix* yang dihasilkan dan menunjukkan akurasi serta perbandingan klasifikasi secara visual. ROC adalah grafik dua dimensi, dengan *false positive* ditunjukkan sebagai garis horizontal dan *true positive* ditunjukkan sebagai garis vertikal. Hasil penelitian menunjukkan bahwa klasifikasi masuk kedalam kategori *good classification* dengan nilai AUC sebesar 0,88.

5. Kesimpulan dan Saran

Berdasarkan hasil eksperimen yang dilakukan untuk sentimen analisis *review* pengguna aplikasi ChatGPT dengan pembagian data 90:10 didapatkan hasil pengujian dari kernel linier dengan nilai *accuracy* 92,23%, *precision* 92,12%, *F1_score* 95,90%, dan AUC 86%. Kernel polinomial dengan nilai *accuracy* 90,78%, *precision* 90,78%, *F1_score* 95,17%, dan AUC 80%. Kernel sigmoid dengan nilai *accuracy* 92,23%, *precision* 92,2%, *F1_score* 95,90%, dan AUC 80%. Parameter terbaik menggunakan kernel RBF didapatkan nilai *accuracy* 92,72%, *precision* 93,43%, dan *F1_score* 96,10%, dan AUC = 88%. Dari hasil penelitian terbukti bahwa kernel pada metode SVM dapat meningkatkan nilai akurasi untuk *review* aplikasi ChatGPT. Berdasarkan pelabelan sentimen menggunakan VADER Lexicon menghasilkan persentase sentimen positif lebih besar daripada sentimen negatif. Hal ini membuktikan bahwa pengguna sangat terbantu dengan adanya aplikasi ChatGPT. Saran untuk penelitian selanjutnya agar dapat membandingkan algoritma SVM dengan beberapa algoritma pengklasifikasi text seperti Naïve Bayes atau K-Nearest Neighbors (KNN).

Referensi

- [1] M. Dowling and B. Lucey, "ChatGPT for (Finance) research: The Bananarama Conjecture," *Financ Res Lett*, vol. 53, May 2023, doi: 10.1016/j.frl.2023.103662.
- [2] Q. Zhong, L. Ding, J. Liu, B. Du, and D. Tao, "Can ChatGPT Understand Too? A Comparative Study on ChatGPT and Fine-tuned BERT," *arxiv.org*, 2022.
- [3] Y. Zhu, P. Zhang, E.-U. Haq, P. Hui, and G. Tyson, "Can ChatGPT Reproduce Human-Generated Labels? A Study of Social Computing Tasks," *Association for Computing Machinery*, vol. 1, no. 1, 2023.
- [4] S. Biswas, "The function of chat GPT in social media: According to chat GPT. Introduction," *SSRN*, pp. 1–6, 2023.
- [5] A. Ahmad and W. Gata, "Sentimen Analisis Masyarakat Indonesia di Twitter Terkait Metaverse dengan Algoritma Support Vector Machine," *Jurnal Teknologi Informasi dan Komunikasi*, vol. 6, no. 4, p. 2022, 2022, doi: 10.35870/jti.
- [6] D. L. Rianti, Y. Umaidah, and A. Voutama, "Tren Marketplace Berdasarkan Klasifikasi Ulasan Pelanggan Menggunakan Perbandingan Kernel," *STRING (Satuan Tulisan Riset dan Inovasi Teknologi)*, vol. 6, no. 1, pp. 98–105, 2021.
- [7] P. H. Prastyo, I. Ardiyanto and R. Hidayat, "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)*, pp. 1-6, 2020, doi: 10.1109/ICDABI51230.2020.9325685.

- [8] R. Indraswari, A. Z. Arifin, and D. Herumurti, "RBF Kernel Optimization Method With Particle Swarm Optimization On SVM Using The Analysis Of Input Data's Movement," *Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information)*, vol. 10, no. 1, pp. 36–42, 2017.
- [9] A. Ramadhanu, R. A. Mahessya, M. R. Zaky, and M. Isra, "Penerapan Teknologi Machine Learning Dengan Metode VADER," *JOISIE Journal Of Information System And Informatics Engineering*, vol. 7, no. 1, pp. 165–173, 2023.
- [10] I. G. S. A. Putu, S. Deviya, Y. Sibaroni, and E. B. Setiawan, "Aspect-Based Sentiment Analysis on iPhone Users on Twitter Using the SVM Method and Optimization of Hyperparameter Tuning," *Jurnal Media Informatika Budidarma*, vol. 7, no. 1, pp. 89–98, 2023.
- [11] M. D. Hendriyanto, A. A. Ridha, and U. Enri, "Analisis Sentimen Ulasan Aplikasi Mola Pada Google Play Store Menggunakan Algoritma Support Vector Machine," *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 5, no. 1, pp. 1–7, 2022.
- [12] S. Ernawati, R. Wati, and N. Nuris, "Support Vector Classification with Hyperparameters for Analysis of Public Sentiment on Data Security in Indonesia," *Jurnal Riset Informatika*, vol. 5, no. 1, pp. 85–92, 2022.
- [13] O. H. Rahman, G. Abdillah, and A. Komarudin, "Klasifikasi Ujaran Kebencian pada Media Sosial Twitter Menggunakan Support Vector Machine," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 1, pp. 17–23, 2021.
- [14] M. Wankhade, A. Chandra, S. Rao, and C. Kulkarni, "A Survey on Sentiment Analysis Methods, Applications, and Challenges," *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5731–5780, 2022.
- [15] A. Haleem, M. Javaid, and R. P. Singh, "An era of ChatGPT as a significant futuristic support tool : A study on features, abilities, and challenges," *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, vol. 2, no. 4, pp. 1–8, 2023.
- [16] Ö. Aydin and E. Karaarslan, "OpenAI ChatGPT Generated Literature Review : Digital Twin in Healthcare," *Emerging Computer Technologies*, vol. 2, pp. 22–31, 2022.
- [17] J. Yin and Q. Li, "A Semismooth Newton Method for Support Vector Classification and Regression," *Computational Optimization and Applications*, vol. 73, pp. 477–508, 2019. DOI: 10.1007/s10589-019-00075-z.
- [18] M. Ahmad, S. Aftab, M. S. Bashir, and N. Hameed, "Sentiment Analysis using SVM : A Systematic Literature Review," (*IJACSA International Journal of Advanced Computer Science and Applications*), vol. 9, no. 2, pp. 182–188, 2018.
- [19] S. I. Nurhafida, F. Sembiring, J. Raya, and C. No, "Analisis Sentimen Aplikasi Novel Online Di Google Play Store Menggunakan Algoritma Support Vector Machine (SVM)," *Jurnal Sains Komputer & Informatika (J-SAKTI)*, vol. 6, no. 1, pp. 317–327, 2022.
- [20] L. Parisi, "m-arcsinh: An Efficient and Reliable Function for SVM and MLP in scikit-learn," 2020, doi: <https://doi.org/10.48550/arXiv.2009.07530>.
- [21] C. Çilgin, M. BAŞ, H. BİLGEHAN, and C. ÜNAL, "Twitter Sentiment Analysis During Covid-19 Outbreak with VADER," *Academic Journal of Information Technology*, vol. 13, no. 49, pp. 73–89, 2022.
- [22] T. Yan, S.-L. Shen, A. Zhou, and X. Chen, "Prediction of geological characteristics from shield operational parameters by integrating grid search and K-fold cross validation into stacking classification algorithm," *Journal of Rock Mechanics and Geotechnical Engineering*, vol. 14, no. 4, pp. 1292–1303, 2022.
- [23] R. G. S. K, A. K. Verma, and S. Radhika, "K-Nearest Neighbors and Grid Search CV Based Real Time Fault Monitoring System for Industries," in *2019 5th International Conference for Convergence in Technology (I2CT)*, 2019, pp. 9–13.
- [24] K. Dataset, "ChatGPT Tweets," *Kaggle*, 2022. [Online]. Available: <https://www.kaggle.com/datasets/furqanrustam118/chatgpt-tweets>. [Accessed: 01-Sep-2023].
- [25] M. Carrington *et al.*, "Deep ROC Analysis and AUC as Balanced Average Accuracy , for Improved Classifier Selection , Audit and Explanation," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, vol. 45, no. 1, pp. 329–341, 2023.
- [26] W. Suharmawan, "Pemanfaatan Chat GPT Dalam Dunia Pendidikan," *Education Journal : Journal Education Research and Development*, vol. 7, no. 2, pp. 158–166, 2023.
- [27] D. M. W. Powers and Ailab, "EVALUATION: FROM PRECISION, RECALL AND F-MEASURE TO ROC, INFORMEDNESS, MARKEDNESS & CORRELATION," 2020, doi: <https://doi.org/10.48550/arXiv.2010.16061>.