

IMPLEMENTASI SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN ROBOT POLISI HUMANOID

Intan Sari B¹, Muh. Rafli Rasyid², Farid Wajidi³

^{1,2,3} Informatika, Teknik, Universitas Sulawesi Barat

e-mail: ¹sari32725@gmail.com, ²mrafli@unsulbar.ac.id, ³faridwajidi@unsulbar.ac.id

Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap kehadiran robot polisi humanoid melalui komentar di platform YouTube. Data diperoleh dari tiga video dengan total 1.978 komentar, kemudian diproses melalui tahapan preprocessing teks, pelabelan sentimen menggunakan metode berbasis leksikon, serta ekstraksi fitur menggunakan TF-IDF. Proses klasifikasi dilakukan dengan algoritma Support Vector Machine (SVM), dan evaluasi dilakukan melalui beberapa skenario pembagian data serta validasi silang. Hasil penelitian menunjukkan bahwa model terbaik menggunakan kernel RBF dengan parameter optimal menghasilkan akurasi 82,96%, presisi 82,61%, dan recall 82,96%. Validasi menggunakan 10-fold cross validation menunjukkan performa yang stabil dan konsisten, dengan rata-rata akurasi 82,24%. Temuan ini memperlihatkan bahwa pendekatan berbasis SVM efektif dalam mengidentifikasi sentimen publik terhadap inovasi teknologi robot kepolisian.

Kata Kunci : Analisis sentimen, Klasifikasi teks, Robot polisi humanoid, Support Vector Machine, YouTube

I. PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah menjadi pendorong utama transformasi besar dalam sistem industri dunia yang dikenal dengan istilah Revolusi Industri 4.0. Era ini memadukan teknologi digital dan sistem fisik melalui otomatisasi cerdas yang didukung oleh konektivitas, seperti artificial intelligence, Internet of Things (IoT), big data, serta teknologi berbasis komputasi awan. Revolusi ini menekankan pada integrasi antara dunia fisik dan digital melalui otomatisasi yang cerdas dan terhubung. Sejumlah teknologi mutakhir menjadi fondasi utama dalam revolusi ini, seperti *Internet of Things* (IoT), kecerdasan buatan (AI), *big data*, komputasi awan, augmented reality, keamanan siber, manufaktur aditif, integrasi sistem, dan simulasi [1]. Diantara berbagai inovasi tersebut Salah satu bentuk inovasi AI yang menarik perhatian publik adalah penerapan robot polisi *humanoid* dalam peran-peran sosial.

Robot humanoid adalah robot yang dirancang menyerupai karakteristik manusia dalam hal struktur dan perilaku. Dengan bantuan sensor dan teknologi kecerdasan buatan, robot ini mampu menirukan ekspresi serta gerakan manusia, sehingga memungkinkan

penggunaannya dalam berbagai bidang seperti pendidikan, industri, dan layanan publik. Penggunaan robot ini sangat beragam, mulai dari asisten pribadi, alat bantu dalam kegiatan riset dan pengembangan, media pembelajaran dalam dunia pendidikan, hingga penerapannya dalam sektor industri manufaktur. Keberadaan robot *humanoid* mencerminkan kemajuan teknologi di bidang robotika dan elektronika, dan menjadi salah satu topik penting yang banyak dibahas dalam konteks Revolusi Industri 4.0 [2]. Beberapa negara, termasuk Indonesia, mulai mengeksplorasi penggunaan robot polisi *humanoid* untuk menunjang operasional kepolisian. Keberadaan robot ini menimbulkan berbagai tanggapan dari masyarakat, baik dalam bentuk dukungan, kekhawatiran, maupun kritik.

Meskipun teknologi robot *humanoid* telah dikembangkan dan diuji di berbagai negara, kajian yang menyoroti tanggapan dan persepsi masyarakat terhadap robot polisi, khususnya di Indonesia, masih terbatas. Hal ini membuka peluang untuk menggali opini publik melalui pendekatan analisis sentimen berbasis teks.

Untuk mengetahui bagaimana penerimaan masyarakat terhadap robot polisi *humanoid*, perlu dilakukan analisis sentimen public dan pemilihan metode yang digunakan untuk klasifikasi sentimen, khususnya dari platform digital seperti media sosial atau forum daring. Analisis ini penting untuk menilai sejauh mana teknologi ini diterima dan apa saja kekhawatiran yang muncul.

Analisis sentimen adalah teknik untuk mengenali dan mengevaluasi tanggapan atau opini masyarakat terhadap suatu permasalahan, tema, maupun produk melalui proses pengumpulan dan pemrosesan data berbasis teks [3].

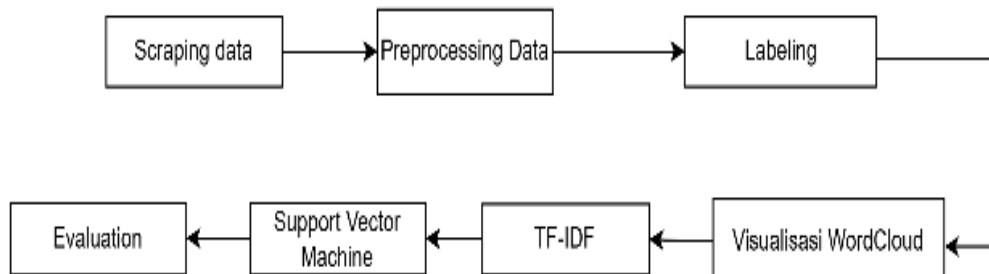
Salah satu algoritma yang terbukti efektif untuk digunakan dalam tugas klasifikasi teks adalah Support Vector Machine (SVM). *Support Vector Machine* (SVM) merupakan algoritma klasifikasi yang mampu membentuk batas pemisah terbaik antar kelas, termasuk dalam situasi data yang tidak linear atau berdimensi tinggi, sehingga sering digunakan dalam tugas-tugas klasifikasi kompleks seperti analisis sentimen. SVM dirancang untuk mengidentifikasi batas pemisah terbaik antar kelas, termasuk dalam kondisi data nonlinier dan berdimensi tinggi, sehingga cocok untuk berbagai kasus klasifikasi teks. [4].

Dengan demikian penelitian ini bertujuan untuk

menganalisis sentimen publik terhadap keberadaan robot polisi *humanoid* dengan menggunakan algoritma SVM,

guna memahami persepsi masyarakat serta memberikan masukan terhadap pengembangan teknologi tersebut ke depan.

II. METODE PENELITIAN



Gambar 1. Alur Penelitian

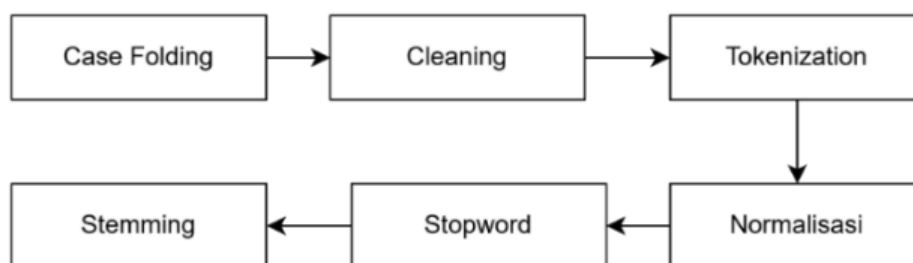
A. Scraping Data

Scraping adalah aktivitas memproses data, khususnya dalam mengambil data dari situs *web* atau media sosial untuk mengungkap pola tersembunyi di dalamnya [5]. Proses pengambilan data dilakukan melalui teknik *scraping* menggunakan *YouTube API* dengan bantuan pustaka *Python*. Data dikumpulkan dari tiga video yang diunggah oleh kanal *YouTube* berbeda, dengan rentang waktu pengambilan mulai dari 27 Juni hingga 4 Juli. Pemilihan rentang waktu ini didasarkan pada momentum menjelang peringatan Hari Bhayangkara pada 1 Juli, di mana robot *humanoid* mulai diperkenalkan dan

dipamerkan kepada publik sebagai bagian dari rangkaian kegiatan menjelang perayaan tersebut. Data yang dikumpulkan mencakup *comment*, dengan jumlah data 1.978.

B. Preprocessing Data

Pra-pemrosesan data teks merupakan tahapan awal yang bertujuan untuk mengonversi teks mentah yang belum terstruktur ke dalam bentuk yang lebih terorganisir, sehingga dapat dianalisis secara komputasi dengan lebih efisien. [6]. Tahapan *preprocessing* data dapat dilihat pada gambar. 2.



Gambar 2. Tahapan *preprocessing*

Case folding merupakan langkah pertama dalam pemrosesan teks yang bertujuan menyeragamkan format penulisan dengan mengonversi semua huruf menjadi huruf kecil agar proses analisis berikutnya lebih mudah dilakukan [7].

Setelah itu lanjut ke tahap *Cleaning*, Pembersihan data (*cleaning*) dilakukan untuk menghilangkan elemen-elemen yang tidak diperlukan seperti simbol, angka, dan karakter khusus, guna meningkatkan kualitas data teks yang akan dianalisis[8].

Langkah berikutnya adalah tokenisasi, yaitu proses membagi teks menjadi unit-unit terkecil seperti kata, setelah terlebih dahulu menghilangkan tanda baca, emoji, dan elemen HTML yang tidak diperlukan. [9].

Normalisasi, normalisasi merupakan langkah krusial dalam analisis sentimen, khususnya saat mengolah data

komentar dari *platform* seperti *YouTube*, karena banyak komentar ditulis dengan bahasa tidak baku [10].

Langkah berikutnya adalah stopwords removal, yakni proses menghilangkan kata-kata yang dianggap tidak memiliki nilai informasi yang signifikan dalam proses analisis atau pengolahan data [11].

Stemming adalah tahapan untuk mengidentifikasi bentuk dasar dari sebuah kata dengan memperhatikan kaidah atau struktur kebahasaan. Tahapan ini dilakukan dengan menghapus berbagai imbuhan, seperti awalan, akhiran, sisipan, maupun kombinasi dari imbuhan tersebut. [12].

C. Labelling

Pelabelan data dilakukan menggunakan VADER, VADER merupakan alat analisis sentimen yang menggunakan pendekatan berbasis leksikon dan aturan,

yang secara khusus dirancang untuk memahami ekspresi sentimen di media sosial. Alat ini menghasilkan skor gabungan yang disebut *compound*, yang merupakan kombinasi dari nilai sentimen positif, negatif, dan netral. Skor *compound* yang bernilai positif mengindikasikan kecenderungan sentimen positif, skor negatif menunjukkan sentimen negatif, sedangkan skor yang mendekati nol merepresentasikan sentimen netral [13].

D. Visualisasi Wordcloud

Word Cloud adalah bentuk visualisasi teks yang menampilkan kumpulan kata, di mana ukuran tiap kata menunjukkan frekuensi kemunculannya dalam suatu dokumen atau dataset. Kata-kata yang lebih sering muncul akan tampak lebih besar, sehingga memudahkan identifikasi kata yang paling dominan dalam konteks analisis [14].

E. TF-IDF

TF-IDF adalah metode yang digunakan untuk menghitung bobot masing-masing kata setelah melewati proses pra-pemrosesan teks [15].

$$TFIDF(t, d, D) = TF(t, d) \times IDF(t, D) \quad (1)$$

Persamaan (1) digunakan untuk menghitung nilai bobot TF-IDF dari setiap token. Dalam hal ini, t menunjukkan token tertentu yang akan diberi bobot, d merujuk pada dokumen tempat token tersebut muncul, sedangkan D merepresentasikan total keseluruhan dokumen dalam korpus yang dianalisis [16].

F. Support Vector Machine

Support Vector Machine (SVM) merupakan algoritma klasifikasi yang mampu membentuk batas pemisah terbaik antar kelas, termasuk dalam situasi data yang tidak linear atau berdimensi tinggi, sehingga sering digunakan dalam tugas-tugas klasifikasi kompleks seperti analisis sentimen [4]. Beberapa tahapan dalam proses implementasi *Support Vector Machine* (SVM) yaitu :

Train-Test Split merupakan langkah awal evaluasi model dengan cara membagi dataset menjadi dua bagian, yakni 70% untuk pelatihan dan 30% untuk pengujian. Model dilatih pada data pelatihan, lalu diuji menggunakan data pengujian guna melihat performa awal klasifikasi SVM. Metrik evaluasi yang digunakan antara lain akurasi, presisi, dan recall.

Hyperparameter Tuning dengan *GridSearchCV*, Setelah melakukan evaluasi awal, tahapan selanjutnya Untuk mendapatkan performa model yang optimal, dilakukan pencarian kombinasi parameter terbaik menggunakan metode *Grid Search* dengan *Cross Validation* (*GridSearchCV*). Parameter yang diuji meliputi nilai C , γ , dan degree sesuai dengan jenis kernel yang digunakan. Proses ini dilakukan pada data latih dengan validasi silang sebanyak 5 lipatan (*5-fold CV*). Model SVM diuji menggunakan tiga jenis kernel, yaitu Linear, *Radial Basis Function* (RBF), dan *Polynomial*. Persamaan dari tiga kernel terdapat pada persamaan (2) kernel linear, (3) kernel RBF, dan (4) kernel *polynomial* [17].

$$K = (X_i, X_j) = (X_i \cdot X_j) \quad (2)$$

$$K = (X_i, X_j) = (X_i \cdot X_j + 1)^p \quad (3)$$

$$K = (X_i, X_j) = e^{-\gamma(X_i \cdot X_j)^2} \quad (4)$$

Untuk setiap jenis kernel, dilakukan proses optimasi *hyperparameter* menggunakan *GridSearchCV*, yaitu metode pencarian parameter terbaik melalui eksplorasi kombinasi nilai parameter tertentu dengan teknik validasi silang (*cross-validation*). Proses ini bertujuan untuk menentukan kernel terbaik.

Evaluasi dengan Tiga Skema Pembagian Data, Setelah menemukan kernel terbaik, Langkah selanjutnya adalah Evaluasi dengan Tiga Skema Pembagian Data Untuk mengetahui stabilitas dan konsistensi kinerja model, dilakukan pengujian dengan tiga variasi rasio pembagian data latih dan data uji, yaitu 80:20, 70:30, dan 90:10. Evaluasi ini membantu mengidentifikasi sensitivitas model terhadap jumlah data latih dan uji yang berbeda.

G. Evaluation

Tahap akhir dalam proses ini adalah evaluasi model. Pada tahap ini, performa model dianalisis dengan menggunakan *confusion matrix* dan sejumlah metrik evaluasi, yaitu akurasi, presisi, *recall*, dan *F1-score*. *F1-score* dipilih sebagai indikator utama karena mampu memberikan gambaran seimbang antara presisi dan *recall*. Selain menggunakan *confusion matrix*, juga menggunakan kurva ROC (*Receiver Operator Characteristic*) dan nilai AUC (*Area Under Curve*).

$$\text{akurasi} = \left(\frac{TP + TN}{TP + TN + FP + FN} \right) * 100 \quad (5)$$

$$\text{presisi} = \left(\frac{TP}{TP + FP} \right) * 100 \quad (6)$$

$$\text{recall} = \left(\frac{TP}{TP + FN} \right) * 100 \quad (7)$$

$$F1_score = 2 * \left(\frac{\text{recall} * \text{precision}}{(\text{recall} + \text{precision})} \right) \quad (8)$$

Persamaan (5), (6), (7) dan (8) digunakan Untuk menghitung akurasi, presisi, *recall*, *F1-score* [18]. Beberapa tahapan dalam proses evaluasi yaitu:

Evaluasi Model Terbaik, Setelah dilakukan optimasi parameter dan pemilihan kernel terbaik, model akhir diuji ulang pada data uji untuk melihat performa riil dari model. Hasil dari uji ini merepresentasikan kinerja akhir model dalam skenario klasifikasi nyata, ditunjukkan dengan akurasi, presisi, *recall* dan Visualisasi digunakan untuk memahami hasil klasifikasi secara lebih dalam. *Confusion matrix* menampilkan jumlah prediksi benar dan salah untuk masing-masing kelas.

Validasi 10-Fold Cross Validation adalah metode validasi yang bertujuan untuk mengevaluasi kemampuan model statistik dalam menggeneralisasi data yang belum pernah digunakan selama proses pelatihan [19]. Teknik validasi silang dilakukan dengan membagi data menjadi 10 bagian, lalu melatih dan menguji model sebanyak 10 kali secara bergantian. Evaluasi ini penting untuk mengukur stabilitas dan kemampuan generalisasi model terhadap data baru yang belum pernah dilihat. Hasilnya menunjukkan seberapa konsisten model dalam berbagai skenario data.

Analisis Visual Evaluasi (*ROC Curve*). Tahapan *ROC Curve* dilakukan sebagai proses evaluasi akhir untuk mengukur kinerja model klasifikasi. Proses dimulai dengan menghitung skor prediksi probabilitas dari model untuk setiap data uji. Selanjutnya, model dievaluasi dengan mengubah *threshold* klasifikasi secara bertahap untuk menghasilkan nilai TPR dan FPR. Nilai-nilai ini kemudian diplot dalam grafik dua dimensi dengan FPR pada sumbu X dan TPR pada sumbu Y.

III. HASIL DAN PEMBAHASAN

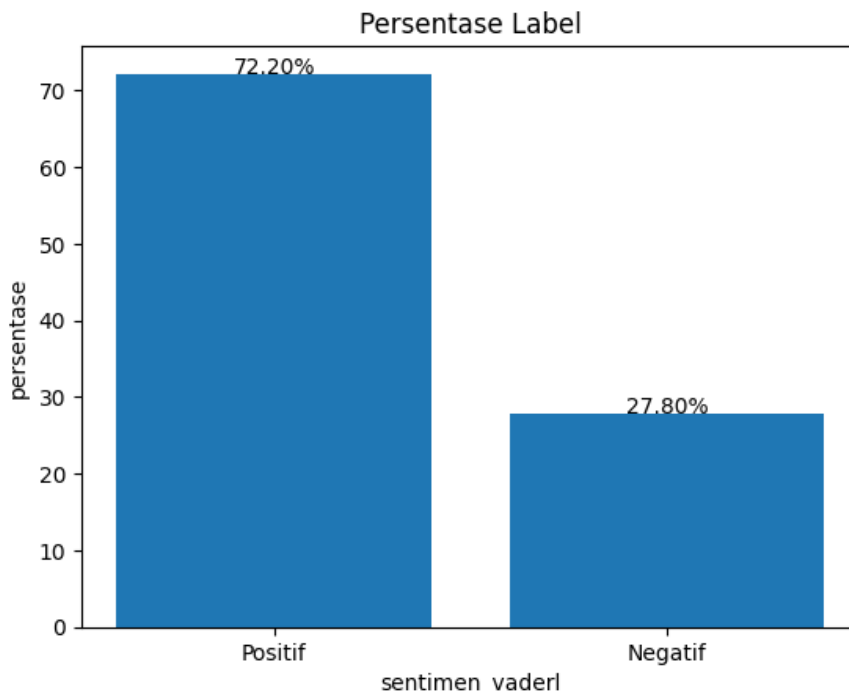
Penelitian ini menganalisis sentiment public dikomentar Youtube terkait Robot polisi *Humanoid*. Data yang diperoleh dari platform Youtube dengan cara *scraping*. Data diambil berupa komentar dari beberapa 3 video dengan channel youtube yang berbeda, total data yaitu 1.978.

Tahap berikutnya adalah melakukan preprocessing data, yang bertujuan untuk mengubah teks mentah yang tidak terstruktur menjadi format yang lebih terorganisir. Hasil dari proses ini disajikan pada Tabel 1.

Tabel 1. Hasil *Preprocessing*

Komentar	Hasil <i>preprocessing</i>
polisi kok dulinan mainan remot	['polisi', 'main', 'main', 'remot']

Setelah melakukan *preprocessing* data, tahap selanjutnya adalah pelabelan sentimen yang dilakukan secara otomatis menggunakan algoritma VADER (*Valence Aware Dictionary and sEntiment Reasoner*). VADER menghasilkan tiga kategori sentimen, yaitu positif, negatif, dan netral. Namun, pada penelitian ini hanya dua kelas yang digunakan, yaitu positif dan negatif, dengan menghapus kelas netral. Penghapusan label netral dilakukan karena sentimen netral dinilai tidak memiliki arah kecenderungan opini yang jelas, bersifat ambigu, dan kurang relevan terhadap fokus penelitian yang bertujuan menganalisis polarisasi opini publik. Selain itu, klasifikasi dua kelas juga memberikan kestabilan dan akurasi model yang lebih baik. Hasil pelabelan menunjukkan bahwa sebanyak 72,20% (805) data termasuk dalam kategori positif, dan 27,80% (310) termasuk dalam kategori negatif, sebagaimana ditunjukkan pada Gambar 3.



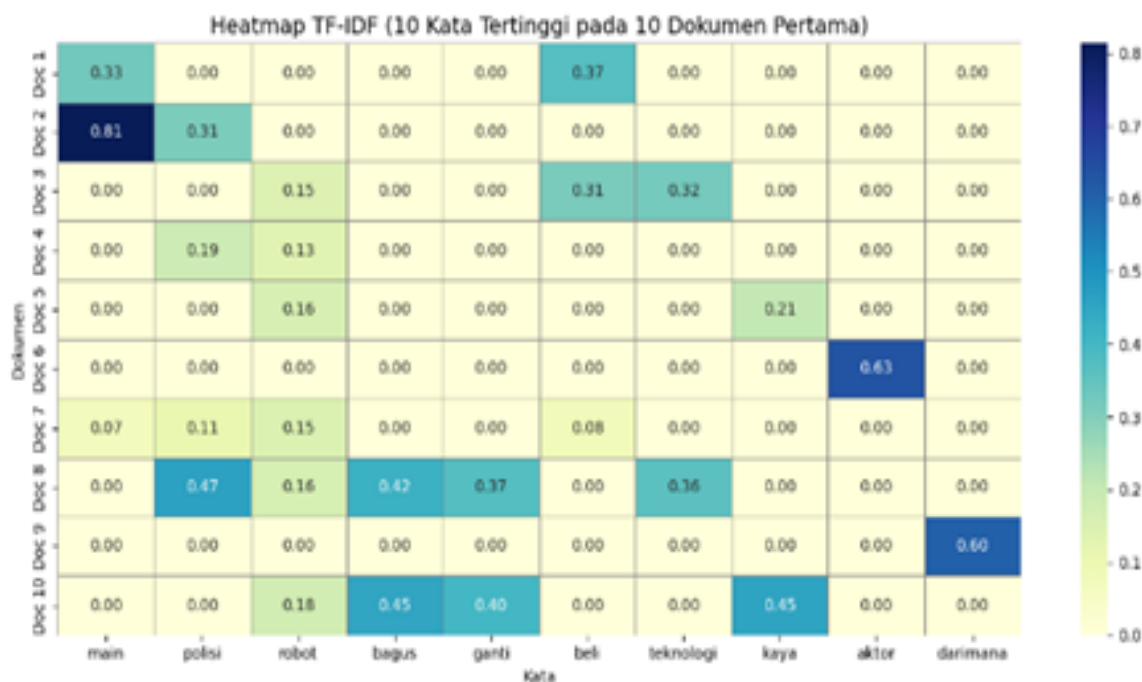
Gambar 3. Persentase hasil labelling

Setelah melakukan pelabelan tahap selanjutnya adalah visualisasi *wordcloud* untuk menampilkan frekuensi kemunculan kata sentiment positif dan negatif. Visualisasinya dapat dilihat pada Gambar 4. Tahapan

selanjutnya yaitu memberikan bobot pada kata-kata yang telah melewati tahap *preprocessing* (TF-IDF).



Gambar 4. *WordCloud* sentiment positif dan negatif



Gambar 5. Hasil TF-IDF

Gambar 5 memperlihatkan visualisasi heatmap dari 10 kata dengan nilai rata-rata TF-IDF tertinggi yang diambil dari 10 dokumen awal dalam korpus. Warna yang lebih gelap menandakan bobot TF-IDF yang lebih tinggi, mengindikasikan bahwa kata tersebut memiliki tingkat kepentingan yang besar dalam dokumen tersebut. Visualisasi ini membantu mengidentifikasi kata-kata kunci yang paling menonjol dan relevan dalam dokumen-dokumen awal, seperti "main", "polisi", "robot", dan "bagus", yang dapat digunakan sebagai fitur penting dalam klasifikasi sentimen menggunakan algoritma SVM. Setelah proses pembobotan kata (TF-IDF) dilanjutkan dengan implementasi algoritma *Support Vector Machine*.

Train-Test Split, pada tahap evaluasi awal, model SVM diuji menggunakan parameter *default* dari pustaka *Scikit-Learn*, yaitu kernel RBF dengan nilai C sebesar 1.0, gamma dengan '*scale*' dan *split* data 70:30. Tujuannya adalah untuk mendapatkan gambaran umum kinerja awal model sebelum dilakukan proses tuning. Pada tahap ini diperoleh hasil pengujian awal dengan model SVM menggunakan kernel RBF dan parameter *default* dari pustaka *Scikit-Learn* menunjukkan hasil akurasi sebesar 80,30%, presisi 84,09%, dan *recall* 90,24%.

Setelah dilakukan optimasi hyperparameter menggunakan metode *GridSearchCV* pada tiga jenis kernel (Linear, RBF, dan *Polynomial*), diperoleh hasil terbaik pada kernel RBF dengan parameter C sebesar 10 dan gamma sebesar 0.1, yang menghasilkan akurasi sebesar 82,96%, presisi 82,61%, dan *recall* 82,96%. Hasil dari pengujian menunjukan kernel RBF lebih unggul dibandingkan Linear dan *polynomial* yang dapat dilihat pada Tabel 2.

Tabel 2. Hasil Uji 3 Kernel

Kernel	Best params	Best score	Akurasi (%)	Presisi (%)	Recall (%)
Linear	{'C': 1}	0,79	77,31	76,31	77,31
RBF	{'C': 10, 'gamma': 0.1}	0,79	82,39	81,64	82,39
Polynomial	{'C': 1, 'degree': 1}	0,79	77,31	76,31	77,31

Selanjutnya, dilakukan evaluasi lanjutan dengan menerapkan tiga rasio pembagian data (70:30, 80:20, dan 90:10). Dari ketiganya, rasio 80:20 memberikan hasil paling optimal, hasil dari tiga rasio pembagian data dapat dilihat pada tabel 3.

Tabel 3. Hasil Uji Kernel Terbaik

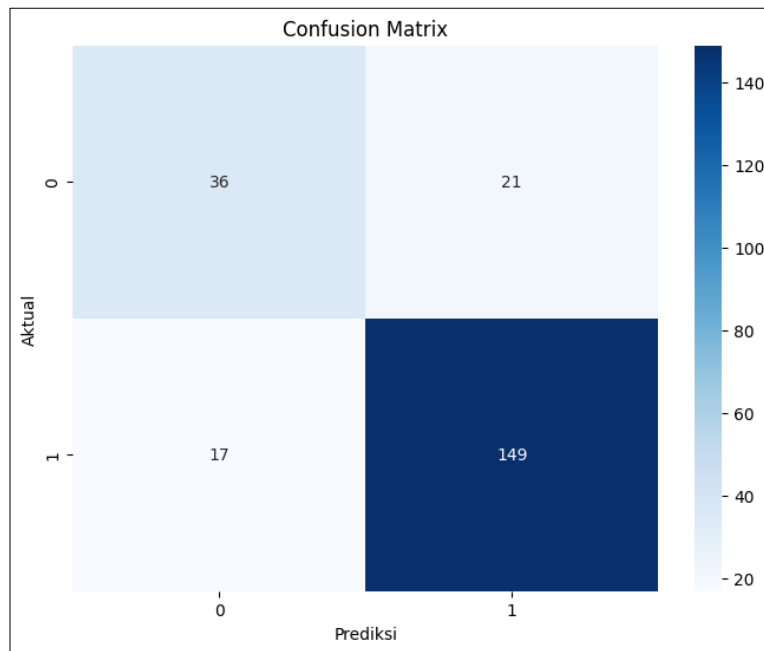
Rasio data	Akurasi (%)	Presisi (%)	Recall (%)
70:30	82,39	81,64	82,39
80:20	82,96	82,61	82,96
90:10	81,25	80,61	81,25

Tahap terakhir adalah evaluasi, pada tahap evaluasi terdapat tiga proses yaitu Evaluasi Model Terbaik, Validasi Generalisasi Model (*10-Fold Cross Validation*) dan Analisis Visual Evaluasi (*ROC Curve*).

Pada evaluasi model terbaik, Setelah dilakukan optimasi parameter dan pemilihan kernel terbaik, model akhir diuji ulang pada data uji untuk melihat performa riil dari model. Model akhir RBF) diuji dengan parameter terbaik yaitu C sebesar 10, gamma sebesar 0,1 dengan split data 80:20. Hasil dari pengujian dapat dilihat pada Tabel 4 dan visualisasi *Confusion matrix* pada Gambar 6 dengan total prediksi benar 185.

Tabel 4. Hasil Evaluasi Model Terbaik

Data latih	Data uji	Akurasi (%)	Presisi (%)	Recall (%)
892	223	82,96	82,61	82,96



Gambar 6. Confusion Matrix

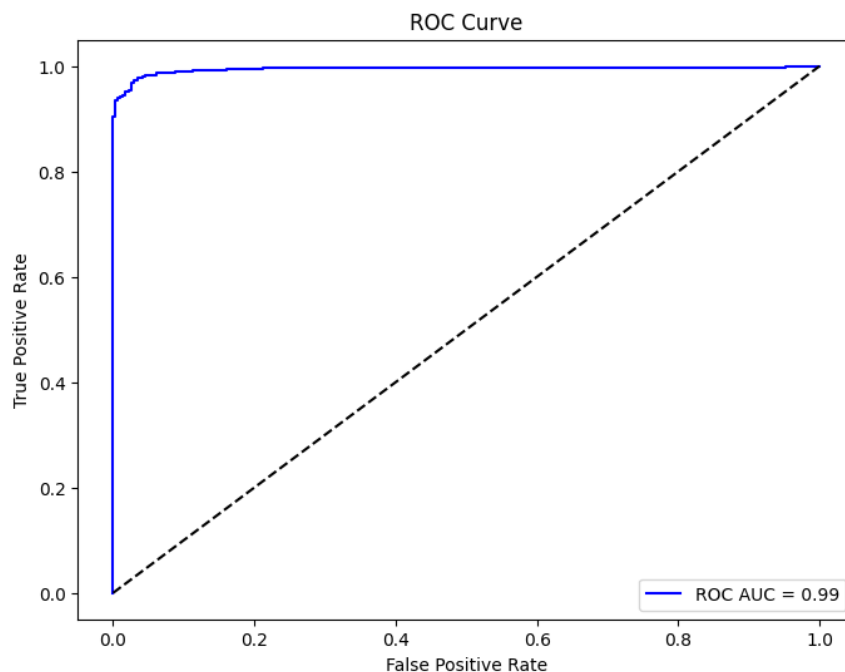
Tahap terakhir yaitu Validasi Generalisasi Model (*10-Fold Cross Validation*) dan Analisis Visual Evaluasi (*ROC Curve*). Hasilnya dapat dilihat pada Tabel 5 dan Gambar 7.

Berdasarkan pengujian menggunakan teknik *10-fold cross validation*, model menunjukkan rata-rata akurasi sebesar 82,24%, dengan presisi mencapai 82,02%, *recall* sebesar 81,44%, serta *F1-score* sebesar 82,24%. Hasil ini mencerminkan performa yang konsisten dan stabil dalam proses klasifikasi. Selain itu, analisis berdasarkan kurva ROC yang ditampilkan pada Gambar 7 Nilai AUC sebesar 0,99 dari hasil kurva ROC memperlihatkan bahwa model memiliki kemampuan klasifikasi yang sangat akurat dalam membedakan dua kelas.. Kurva ROC yang

mendekati sudut kiri atas grafik merupakan indikasi dari kualitas klasifikasi yang tinggi.

Tabel 5. 10-Fold Cross Validation

Fold	Akurasi (%)	Presisi (%)	Recall (%)	F1-Score (%)
1	83,04	82,21	83,04	82,41
2	81,25	80,42	81,25	80,45
3	82,14	81,86	82,14	80,71
4	85,71	85,41	85,71	85,43
5	80,36	80,30	80,36	78,62
6	81,08	80,35	81,08	79,65
7	80,18	79,48	80,18	79,71
8	81,08	82,08	81,08	79,64
9	81,98	81,98	81,98	81,98
10	85,59	86,09	85,59	85,80



Gambar 7. ROC Curve

Evaluasi menggunakan skema *10-fold cross validation* menunjukkan konsistensi model, dengan rata-rata akurasi 82,24%, presisi 82,02%, *recall* 81,24%, dan *F1-score* 82,24%. Nilai AUC sebesar 0,99 dari kurva ROC

menunjukkan bahwa model memiliki kemampuan klasifikasi yang sangat presisi dalam membedakan sentimen positif dan negatif.

IV. KESIMPULAN DAN SARAN

Penelitian ini bertujuan untuk mengevaluasi opini publik terhadap kehadiran robot polisi *humanoid* berdasarkan komentar di *platform YouTube*, dengan menerapkan pendekatan klasifikasi menggunakan algoritma *Support Vector Machine (SVM)*. Pelabelan sentimen dengan algoritma *VADER* menunjukkan bahwa sebanyak 72,20% komentar bersifat positif, dan sisanya 27,80% bernuansa negatif. Hasil ini diperkuat oleh visualisasi *Word Cloud* yang menampilkan kata-kata kunci paling sering muncul dalam masing-masing kategori sentimen.

Model *SVM* awal yang diuji dengan parameter bawaan memberikan hasil performa yang cukup tinggi, dengan akurasi 80,30%, presisi 84,09%, dan *recall* 90,24%. Setelah dilakukan proses *tuning hyperparameter* melalui *GridSearchCV* terhadap tiga kernel yang berbeda (*Linear*, *RBF*, dan *Polynomial*), diperoleh bahwa kernel *RBF* dengan *C* sebesar 10 dan *gamma* sebesar 0.1 menunjukkan performa terbaik. Pengujian model dengan konfigurasi ini pada skenario pembagian data 80:20 menghasilkan akurasi 82,96%, presisi 82,61%, dan *recall* 82,96%.

Validasi lanjutan menggunakan 10-fold *cross validation* mengindikasikan bahwa model memiliki performa yang konsisten dan andal, dengan rata-rata akurasi 82,24%, presisi 82,02%, *recall* 81,24%, dan *F1-score* 82,24%. Selain itu, nilai *AUC* sebesar 0,99 dari analisis kurva *ROC* menunjukkan kemampuan klasifikasi yang sangat akurat dalam membedakan sentimen positif dan negatif.

Secara keseluruhan, hasil penelitian ini menegaskan bahwa algoritma *SVM* dengan kernel *RBF* yang telah dioptimasi dapat menjadi pendekatan yang efektif untuk mengklasifikasikan sentimen dalam komentar pengguna. Temuan ini diharapkan dapat memberikan kontribusi terhadap pemahaman persepsi masyarakat terhadap inovasi teknologi robot polisi *humanoid*, sekaligus menjadi referensi dalam perumusan kebijakan dan strategi implementasi teknologi serupa di masa depan.

DAFTAR PUSTAKA

- [1] M. R. Y. Fahdillah, M. Kadar, I. Hassandi, and M. R. "Implementasi Transformasi Digital dan Kecerdasan Buatan Sebagai Inovasi Untuk UMKM pada Era Revolusi Industri 4.0," *Jurnal Ilmiah Manajemen dan Kewirausahaan (JUMANAGE)*, vol. 3, no. 1, Jan. 2024, doi: 10.33998/jumanage.2024.3.1.1552.
- [2] Y. Axel Radityoadji, M. Syarifuddin Zuhrie, L. Rahmawati, and A. Wiyono, "Pengembangan Sistem Kendali Pada Purwarupa Robot Humanoid Untuk Autism Spectrum Disorder Berbasis IoT," 2025. doi: <https://doi.org/10.26740/jte.v14n1.p34-42>.
- [3] Y. Akbar and T. Sugiharto, "Analisis Sentimen Pengguna Twitter di Indonesia Terhadap ChatGPT Menggunakan Algoritma C4.5 dan Naïve Bayes (Yuma Akbar 1*, Tri Sugiharto 2) Analisis Sentimen Pengguna Twitter di Indonesia Terhadap ChatGPT Menggunakan Algoritma C4.5 dan Naïve Bayes," *Jurnal Sains dan Teknologi*, vol. 5, no. 1, pp. 115–122, 2023, doi: 10.55338/saintek.v4i3.1368.
- [4] S. P. Azzahra, Y. A. Apriyanto, and A. Wijaya, "Analisis Sentimen Ulasan Aplikasi Deepl Pada Google Play Dengan Metode Support Vector Machine (SVM)," *Jurnal Sistem Informasi (Jusin)*, vol. 4, no. 2, pp. 59–66, 2023, doi: <https://doi.org/10.32546/jusin.v4i2.2368>.
- [5] Yuniar Eka, Safiroh Dwi, and Wahyuningsih Dian, "revisi-implementasi-scraping-data-untuk-sentimental-pengguna-dompot-digital-dengan-menggunakan-algoritma-machine-learning+35-42," *Jurnal Janitra Informatika dan Sistem Informasi*, vol. 2, Apr. 2022, doi: 10.25008/janitra.v2i1.145.
- [6] I. Verawati and B. S. Audit, "Algoritma Naïve Bayes Classifier Untuk Analisis Sentiment Pengguna Twitter Terhadap Provider By.u," *Jurnal Media Informatika Budidarma*, vol. 6, no. 3, p. 1411, Jul. 2022, doi: 10.30865/mib.v6i3.4132.
- [7] D. Rifaldi, Abdul Fadlil, and Herman, "Teknik Preprocessing Pada Text Mining Menggunakan Data Tweet 'Mental Health,'" *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 3, no. 2, pp. 161–171, Apr. 2023, doi: 10.51454/decode.v3i2.131.
- [8] M. Saputra and M. Iqbal, "Analisis Sentimen Masyarakat Terhadap PON XII Aceh-Sumut Menggunakan Algoritma Naïve Bayes," *SIMKOM*, vol. 10, no. 1, pp. 39–48, Jan. 2025, doi: 10.51717/simkom.v10i1.679.
- [9] M. R. Fahlevvi, "Analisis Sentimen Terhadap Ulasan Aplikasi Pejabat Pengelola Informasi dan Dokumentasi Kementerian Dalam Negeri Republik Indonesia di Google Playstore Menggunakan Metode Support Vector Machine," *Jurnal Teknologi dan Komunikasi Pemerintahan*, vol. 4, no. 1, pp. 1–13, 2022, [Online]. Available: <http://ejournal.ipdn.ac.id/JTKP>,
- [10] P. Ganda Dewata, A. Rizky, and H. Irsyad, "Analisis Sentimen Terhadap Boikot Produk Israel Menggunakan Algoritma Naïve Bayes Dan SMOTE," 2024. Accessed: Jul. 11, 2025. [Online]. Available: <https://rekayasainformatika.com/index.php/JREIN/article/view/1>
- [11] Abd. A. Syam, G. Hardy M, A. Salim, D. F. Surianto, and M. Fajar B, "Analisis Teknik Preprocessing Pada Sentimen Masyarakat Terkait Konflik Israel-Palestina Menggunakan Support Vector Machine," *JIP (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 3, pp. 1464–1472, Aug. 2024, doi: 10.29100/jipi.v9i3.5527.
- [12] M. E. Apriyani, R. Renaldi, T. A. Cinderatama, J. T. Informasi, and P. N. Malang, "JIP (Jurnal Informatika Polinema) Analisis Sentimen Berita Hoax Menggunakan Naïve Bayes," *JIP (Jurnal Informatika Polinema)*, Nov. 2024, doi: <http://dx.doi.org/10.33795/jip.v11i1.4623>.
- [13] S. Ernawati and R. Wati, "Evaluasi Performa Kernel SVM dalam Analisis Sentimen Review Aplikasi ChatGPT Menggunakan Hyperparameter dan VADER Lexicon," 2024.
- [14] J. A. Wibowo, V. C. Mawardi, and T. Sutrisno, "Visualisasi Word Cloud Hasil Analisis Sentimen Berbasis Fitur Layanan Aplikasi Gojek Dengan Support Vector Machine," *Jurnal Serina Sains, Teknik dan Kedokteran*, vol. 2, no. 1, pp. 61–70, Mar. 2024, doi: 10.24912/jsstk.v2i1.32058.
- [15] S. F. Huwaida, R. Kusumawati, and B. Isnaini, "Analisis Sentimen Komentar YouTube terhadap Pemindahan Ibu Kota Negara Menggunakan Metode Naïve Bayes," *Jambura Journal of Informatics*, vol. 6, no. 1, pp. 26–39, Apr. 2024, doi: 10.37905/jji.v6i1.24718.
- [16] Robert Antonius, A. R. Zulkarnain, and H. Irsyad, "Pendekatan TF-IDF, SMOTE, dan SVM dalam Klasifikasi Sentimen Masyarakat terhadap Pemblokiran Judi Online," *Buletin Ilmiah Informatika Teknologi*, vol. 2, no. 3, pp. 115–122, Jun. 2024, doi: 10.58369/biit.v2i3.65.
- [17] T. Meisya, P. Aulia, N. Arifin, and R. Mayasari, "Perbandingan Kernel Support Vector Machine (Svm) Dalam Penerapan Analisis Sentimen Vaksinasi Covid-19," *SINTECH Journal*, vol. 4, Oct. 2021, doi: <https://doi.org/10.31598/sintechjournal.v4i2.762>.
- [18] J. E. Br Sinulingga and H. C. K. Sitorus, "Analisis Sentimen Opini Masyarakat terhadap Film Horor Indonesia Menggunakan Metode SVM dan TF-IDF," *Jurnal Manajemen Informatika (JAMIKA)*, vol. 14, no. 1, pp. 42–53, Feb. 2024, doi: 10.34010/jamika.v14i1.11946.
- [19] D. E. Yanti, L. Framesti, and A. Desiani, "JIP (Jurnal Informatika Polinema) Perbandingan Algoritma C4.5 Dan Svm Dalam Klasifikasi Penyakit Anemia," *JIP (Jurnal Informatika Polinema)*, vol. 9, Aug. 2023, doi: <https://doi.org/10.33795/jip.v9i4.1381>.