

Terbit online pada laman : <http://teknosi.fti.unand.ac.id/>

Jurnal Nasional Teknologi dan Sistem Informasi

| ISSN (Print) 2460-3465 | ISSN (Online) 2476-8812 |



Perbandingan Algoritma Support Vector Machine, Decision Tree, dan Logistic Regresion Pada Analisis Sentimen Ulasan Aplikasi Netflix

Nevita Cahaya Ramadani ^{a,*}, Imam Tahyudin ^b, Azhari Shouni Barkah ^c

Universitas Amikom Purwokerto, Jl. Let. Jend. Pol. Soemarto (depan SPN) Purwokerto, Indonesia

INFORMASI ARTIKEL

Sejarah Artikel:

Diterima Redaksi: 11 Juli 2024

Revisi Akhir: 30 Agustus 2024

Diterbitkan Online: 31 Agustus 2024

KATA KUNCI

Support Vector Machine,
Decision Tree,
Logistic Regression,
Sentimen Analysis

KORESPONDENSI

E-mail: *nevitacahayar1@gmail.com

ABSTRACT

Penelitian ini bertujuan untuk menganalisis sentimen terhadap ulasan pengguna Netflix menggunakan algoritma machine learning seperti Support Vector Machine (SVM), Decision Tree dan Logistic Regression. Dataset yang terdiri dari 3000 ulasan pengguna diambil dari Google Play Store dan melalui proses preprocessing teks yang mencakup penghapusan karakter, tokenisasi, penghapusan stopwords, stemming, serta penyaringan token pendek. Metode TF-IDF digunakan untuk ekstraksi dan pembobotan fitur dalam analisis. Evaluasi hasil menunjukkan bahwa SVM secara konsisten memberikan akurasi yang lebih tinggi dibandingkan Decision Tree dan Logistic Regression dalam klasifikasi sentimen, dengan SVM mencapai akurasi rata-rata 88.18% dan puncak tertinggi 92.69% dalam K-Fold Cross Validation. Implikasi praktis dari penelitian ini adalah Netflix dapat memanfaatkan analisis sentimen untuk meningkatkan pengalaman pengguna dan pengelolaan layanan lebih baik.

1. PENDAHULUAN

Perkembangan teknologi informasi telah mengubah secara signifikan cara manusia mengonsumsi konten hiburan, terutama dengan adopsi layanan streaming video yang semakin populer seperti Netflix. Sebagai platform global yang menawarkan berbagai macam film, serial televisi, dan konten orisinal, Netflix telah menjadi bagian penting dari gaya hidup modern di berbagai belahan dunia. Pengguna Netflix tidak hanya menikmati akses ke ribuan judul hiburan, tetapi juga memberikan umpan balik melalui ulasan yang mereka tinggalkan di platform ini. Ulasan pengguna ini bukan hanya mencerminkan preferensi individu, tetapi juga dapat memberikan wawasan berharga bagi Netflix untuk memahami kepuasan pengguna, kecenderungan dalam preferensi konten, serta masalah yang mungkin dihadapi oleh pengguna dalam penggunaan aplikasi.

Dalam konteks ini, penelitian ini bertujuan untuk menjawab beberapa pertanyaan kunci yang relevan dalam analisis sentimen ulasan pengguna Netflix. Pertama, penelitian akan mengidentifikasi dan menganalisis persebaran sentimen positif, negatif, dan netral dalam ulasan yang diberikan oleh pengguna Netflix. Hal ini penting untuk memahami gambaran keseluruhan dari respons pengguna terhadap layanan Netflix secara keseluruhan [1]. Kedua, penelitian ini akan mengevaluasi dan membandingkan performa dua algoritma klasifikasi yang umum digunakan, yaitu Support Vector Machine (SVM), Decision Tree, dan Logistic Regression dalam mengklasifikasikan sentimen dari ulasan pengguna Netflix. Evaluasi ini akan memberikan pandangan tentang mana algoritma yang lebih efektif dalam konteks klasifikasi sentimen untuk data ulasan Netflix.

Penelitian oleh [2], dengan judul "Komparasi Algoritma Naive Bayes, Support Vector Machine, dan Logistic Regression Pada Analisis Sentimen Pengguna Aplikasi Transportasi Online". Tujuan penelitian ini untuk mengetahui perbandingan

keakuratan antara algoritma Support Vector Machine, Naive Bayes, dan Logistic Regression dengan melakukan analisis sentimen pada data ulasan aplikasi Grab. Hasil pengujian komparasi ditemukan bahwa algoritma Naive bayes memiliki kinerja terbaik dibandingkan algoritma klasifikasi lainnya dengan akurasi yang di dapat algoritma Naive bayes sebesar 88.5%, sedangkan algoritma Support Vector Machine memiliki akurasi terendah dengan akurasi sebesar 85.5%. Sehingga dapat disimpulkan bahwa algoritma Naive bayes memiliki nilai yang lebih baik dibandingkan algoritma Logistic Regression dan Support Vector Machine.

Penelitian lainnya oleh [3] dengan judul “Analisis Sentimen Opini Terhadap Vaksin Covid-19 Pada Media Sosial Twitter Menggunakan Naive Bayes dan Decision Tree”. Tujuan penelitian ini untuk membantu masyarakat dalam memberikan opini terhadap vaksin Covid-19 menggunakan Naive Bayes dan Decision Tree. Hasil yang di dapat dari analisis menunjukkan bahwa pada umumnya masyarakat memeberikan tanggapan positif terhadap kebijakan vaksinasi yang dilakukan pemerintah indonesia.

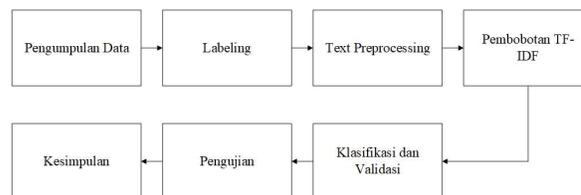
Peelitian lainnya oleh [4] dengan judul “Analisis Klasifikasi Sentimen Terhadap Isu Kebocoran Data Kartu Identitas Ponsel di Twitter”. Tujuannya adalah untuk mencari tahu sebaran kata dan analisis klasifikasi sentimen dari opini masyarakat di Twitter terkait dengan isu tersebut. Hasil pembangunan model menunjukkan algoritma Support-Vector Machine memiliki performa terbaik dengan nilai f1-score 0.81, dilanjutkan dengan Random Forest sebesar 0.78, IndoBERT sebesar 0.76, dan Logistic Regression sebesar 0.74. Ketidakseimbangan kelas dan kurangnya data latih membuat performa IndoBERT sebagai salah satu state-of-the-art dalam NLP memiliki performa yang rendah dibandingkan algoritma lainnya.

Tujuan utama dari penelitian ini adalah untuk memberikan wawasan yang mendalam kepada Netflix tentang bagaimana pengguna mengevaluasi layanan mereka berdasarkan ulasan yang mereka tinggalkan. Dengan memahami sentimen ini, Netflix dapat mengidentifikasi area yang perlu diperbaiki atau ditingkatkan dalam layanan mereka, serta mempertahankan aspek positif yang disukai oleh pengguna. Selain itu, penelitian ini juga bertujuan untuk mengeksplorasi bagaimana teknik preprocessing teks, seperti penghapusan karakter, tokenisasi, dan stemming, dapat mempengaruhi akurasi klasifikasi sentimen. Hal ini penting karena proses preprocessing teks yang efektif dapat meningkatkan kualitas data input untuk algoritma machine learning, yang pada gilirannya dapat meningkatkan performa dan kehandalan sistem klasifikasi sentimen.

Dengan demikian, hasil dari penelitian ini diharapkan dapat memberikan kontribusi yang signifikan baik secara praktis maupun teoritis. Secara praktis, Netflix dapat menggunakan hasil analisis sentimen ini untuk menginformasikan keputusan bisnis mereka, seperti perbaikan produk, pengembangan konten baru, atau peningkatan layanan pelanggan. Secara teoritis, penelitian ini dapat menambah wawasan dalam bidang analisis sentimen dan penerapan machine learning dalam domain klasifikasi teks, memberikan landasan untuk penelitian lebih lanjut dalam analisis

ulasan pengguna dan aplikasi lain dari teknik machine learning dalam industri hiburan digital.

2. METODE



Gambar 1. Metode Penelitian

2.1. Pengumpulan Data

Penelitian ini menggunakan dataset aplikasi netflix yang diambil dari komentar aplikasi netflix di google playstore menggunakan scrapping python google playstore. Dataset yang diambil berjumlah 3000 dataset. Berikut adalah dataset netflix.csv.

	content	score
0	Nonton film di tv ngga full screen layar, jadi...	1
1	Selamat sore Netflix ini pengguna netflix di s...	5
2		5
3		5
4	netflix masih jd apk kecintaan buat nonton fil...	5

Gambar 2. Dataset Netflix

2.2. Labeling

labeling pada penelitian ini digunakan untuk mengelompokan nilai pada kolom score menjadi positive (1), negative (-1), dan netral (0). Jika nilai score >3 akan dikelompokan kedalam sentimen positive (1), jika nilai score >3 akan dikelompokan kedalam sentimen negative (-1), dan jika nilai score 3 akan dikempokan menjadi netral (0) [5].

2.3. Text Preprocessing

Pada penelitian ini melakukan preprocessing text berupa penghapusan karakter, tokenize, stopword, stemming, penghapusan karakter kurang dari 4 kata. berikut adalah penjelasannya:

2.3.1. Penghapusan Karakter

Pada tahap ini dilakukan penghapusan karakter, angka, tanda baca, dan lain sebagainya [6].

Tabel 1. Penghapusan Karakter

Asli	Preprocessing
sangat seru cocok untuk pencinta drakor/anime	sangat seru cocok untuk pencinta drakoranime

2.3.2. Tokenize

Tokenisasi adalah proses memecah teks menjadi bagian-bagian yang lebih kecil [7].

Tabel 2. Tokenize

Preprocessing	Tokenize
sangat seru cocok untuk pencinta drakor/anime	['sangat', 'seru', 'cocok', 'untuk', 'pencinta', 'drakor', 'anime']

2.3.3. Stopword

Pada tahap ini dilakukan untuk menghilangkan kata-kata umum yang tidak memberikan banyak informasi kontekstual [8].

Tabel 3. Stopword

Tokenize	Stopword
['sangat', 'seru', 'cocok', 'untuk', 'pencinta', 'drakor', 'anime']	['seru', 'cocok', 'pencinta', 'drakor', 'anime']

2.3.4. Stemming

Pada tahap ini dilakukan untuk menyederhanakan berbagai bentuk kata menjadi satu bentuk dasar, sehingga memudahkan analisis teks [9].

Tabel 4. Stemming

Stopword	Stemming
['seru', 'cocok', 'pencinta', 'drakor', 'anime']	['seru', 'cocok', 'cinta', 'drakor', 'anime']

2.3.5. Penghapusan Karakter Kurang Dari 4 Kata

Pada tahapan ini dilakukan penghapusan token yang kurang dari 4 huruf digunakan untuk mengurangi noise dalam teks dan fokus pada kata-kata yang lebih bermakna [10].

Tabel 5. Hapus Kata

Stemming	Hapus Kata
['seru', 'cocok', 'cinta', 'drakor', 'anime']	seru cocok cinta drakor anime

Pada tabel 5 tidak berubah itu dikarenakan tidak ada kata yang kurang dari 4 huruf.

2.4. Pembobotan TF-IDF

Pada tahap ini dilakukan pembobotan TF-IDF untuk mengukur relevansi kata dalam konteks dokumen tertentu [11]. Pembobotan TF-IDF ini dilakukan dengan max_features = 2500, berikut adalah daftar TFIDF.

Tabel 6. TF-IDF

term	tfidf
bagus	242.960541
aplikasi	133.937176
good	116.093599
sangat	115.471.402
film	106.958993
mantap	103.364437
bisa	91.602613
keren	76.110831
netflix	67.998000
nonton	63.539290

Setelah itu dilakukan proses pengujian algoritma SVM dan Decision Tree yang akan dilakukan pada bab 3.

3. HASIL

Pada tahap ini, dilakukan pengujian terhadap algoritma Support Vector Machine, Decision Tree dan Logistic Regression, serta pengujian validitas data menggunakan Cross Validation.

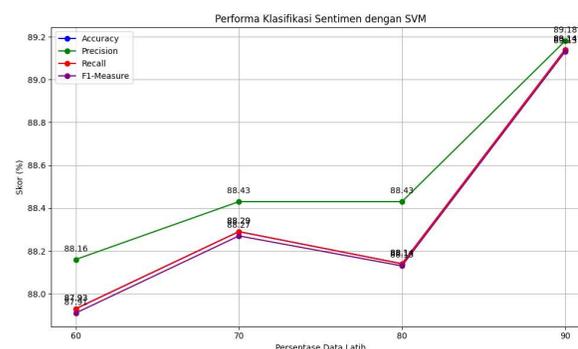
3.1. Pengujian Support Vector Machine

SVM (Support Vector Machine) adalah algoritma klasifikasi yang mencari hyperplane terbaik untuk memisahkan data ke dalam kelas-kelas yang berbeda dengan memaksimalkan margin [2].

Tabel 7. Klasifikasi Sentimen Support Vector Machine

Klasifikasi Sentimen Support Vector Machine				
DL:DU	Accuracy	Precision	Recall	F1-Measure
90%:10%	89.14%	89.18%	89.14%	89.13%
80%:20%	88.14%	88.43%	88.14%	88.13%
70%:30%	88.29%	88.43%	88.29%	88.27%
60%:40%	87.93%	88.16%	87.93%	87.91%

Pengujian Support Vector Machine menunjukkan performa terbaik dengan menggunakan 90% data untuk pelatihan dan 10% data untuk pengujian, mencapai akurasi sebesar 89.14%, serta nilai precision, recall, dan F-Measure masing-masing sekitar 89.18%, 89.14%, dan 89.13%. Penelitian sebelumnya [12] menyarankan bahwa peningkatan jumlah data pelatihan cenderung meningkatkan kinerja model klasifikasi teks. Hasil pengujian yang menggambarkan dampak variasi jumlah data pelatihan dan pengujian dapat dilihat secara rinci pada Gambar 3.



Gambar 3. Grafik Uji Data Latih dan Data Uji Support Vector Machine

3.2. Pengujian Decision Tree

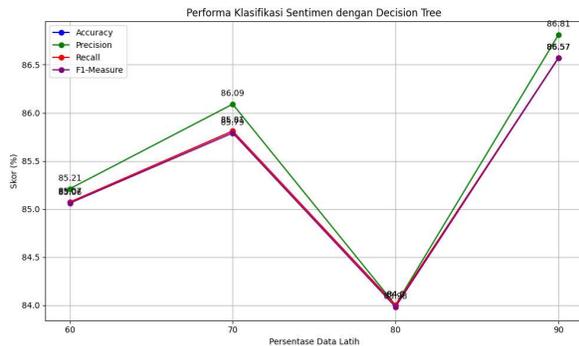
Decision Tree merupakan algoritma yang digunakan untuk memisahkan data ke dalam kelas-kelas berdasarkan serangkaian aturan keputusan yang hierarkis [13].

Tabel 8. Klasifikasi Sentimen Decision Tree

Klasifikasi Sentimen Decision Tree				
DL:DU	Accuracy	Precision	Recall	F1-Measure
90%:10%	86.57%	86.81%	86.57%	86.57%
80%:20%	84.00%	84.00%	84.00%	83.98%
70%:30%	85.81%	86.09%	85.81%	85.79%
60%:40%	85.07%	85.21%	85.07%	85.06%

Pengujian Decision Tree menunjukkan performa paling baik dengan menggunakan 90% data untuk pelatihan dan 10% data untuk pengujian, mencapai akurasi 86.57% serta nilai precision, recall, dan F-Measure sekitar 86.81%, 86.57%, dan 86.57%.

Temuan ini konsisten dengan hasil dari classification report seperti yang dijelaskan oleh [12] bahwa meningkatkan jumlah data pelatihan cenderung meningkatkan performa model klasifikasi teks. Gambar 4 menampilkan secara rinci dampak variasi jumlah data latihan dan pengujian.



Gambar 4. Uji Data Latih dan Data Uji Decision Tree

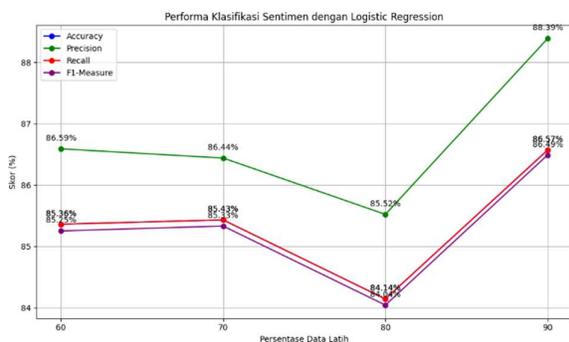
3.3. Pengujian Logistic Regression

Logistic Regression merupakan salah satu metode yang umum digunakan untuk memprediksi kelas atau label dari data berdasarkan input fitur yang diberikan [14].

Tabel 9. Klasifikasi Sentimen Logistic Regression

Klasifikasi Sentimen Logistic Regression				
DL:DU	Accuracy	Precision	Recall	F1-Measure
90%:10%	86.57%	88.39%	86.57%	86.49%
80%:20%	84.14%	85.52%	84.14%	84.04%
70%:30%	85.43%	86.44%	85.43%	85.33%
60%:40%	85.36%	86.59%	85.36%	85.25%

Pengujian Logistic Regression menunjukkan performa paling baik dengan menggunakan 90% data untuk pelatihan dan 10% data untuk pengujian, mencapai akurasi 86.57% serta nilai precision, recall, dan F-Measure sekitar 88.39%, 86.57%, dan 86.49%. Temuan ini konsisten dengan hasil dari classification report seperti yang dijelaskan oleh [15] bahwa meningkatkan jumlah data pelatihan cenderung meningkatkan performa model klasifikasi teks. Gambar 5 menampilkan secara rinci dampak variasi jumlah data latihan dan pengujian.



Gambar 5. Uji Data Latih dan Data Uji Logistic Regression

3.4. Pengujian Cross Validation

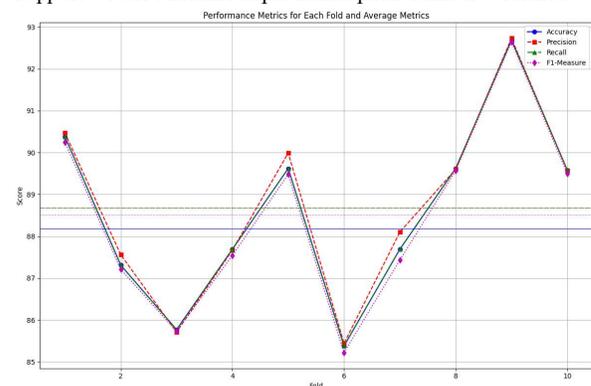
Pengujian Cross Validation pada algoritma Support Vector Machine dilakukan untuk memvalidasi data berdasarkan hasil

terbaik dari pengujian sebelumnya [16], yang disajikan dalam Tabel 10.

Tabel 10. Pengujian Cross Validation Support Vector Machine

Fold Ke-	Accuracy	Precision	Recall	F1-Measure
1	90.38%	90.47%	90.38%	90.25%
2	87.31%	87.56%	87.31%	87.21%
3	85.77%	85.71%	85.77%	85.73%
4	87.69%	87.67%	87.69%	87.54%
5	89.62%	89.99%	89.62%	89.48%
6	85.38%	85.44%	85.38%	85.22%
7	87.69%	88.10%	87.69%	87.43%
8	89.62%	89.60%	89.62%	89.57%
9	92.69%	92.73%	92.69%	92.64%
10	89.58%	89.57%	89.58%	89.50%
Rerata	88.18%	88.68%	88.68%	88.51%

Berdasarkan Tabel 10, terlihat bahwa performa tertinggi dicapai pada iterasi kesembilan dengan akurasi 92.69%, presisi 92.73%, recall 92.69%, dan F-measure 92.64%. Detail hasil pengujian K-Fold Cross Validation menggunakan algoritma Support Vector Machine dapat dilihat pada Gambar 5 berikut.



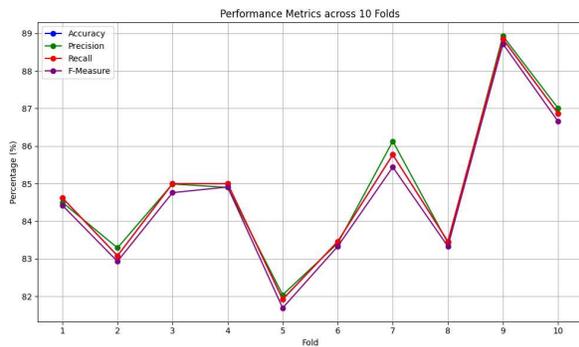
Gambar 6. Grafik Cross Validation Support Vector Machine

Berikutnya adalah pengujian cross validation Decision tree, seperti pada tabel 11.

Tabel 11. Pengujian Cross Validation Decision Tree

Fold Ke-	Accuracy	Precision	Recall	F1-Measure
1	84.62%	84.50%	84.62%	84.42%
2	83.08%	83.29%	83.08%	82.93%
3	85.00%	84.99%	85.00%	84.76%
4	85.00%	84.90%	85.00%	84.91%
5	81.92%	82.04%	81.92%	81.69%
6	83.46%	83.41%	83.46%	83.33%
7	85.77%	86.13%	85.77%	85.44%
8	83.46%	83.42%	83.46%	83.33%
9	88.85%	88.93%	88.85%	88.72%
10	86.87%	87.01%	86.87%	86.66%
Rerata	84.70	84.86%	84.76%	84.52%

Berdasarkan Tabel 11, hasil pengujian K-Fold Cross Validation pada algoritma Decision Tree menunjukkan bahwa performa tertinggi tercapai pada iterasi kesembilan dengan nilai akurasi 88.85%, presisi 88.93%, recall 88.85%, dan F-measure 88.72%. Detail hasil pengujian K-Fold Cross Validation pada algoritma Decision Tree dapat dilihat pada Gambar 7 berikut.



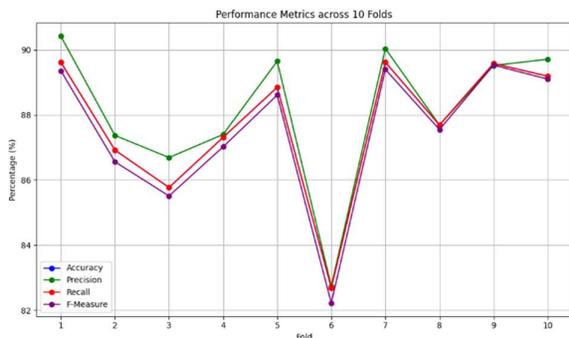
Gambar 7. Grafik Cross Validation Decision Tree

Berikutnya adalah pengujian cross validation Logistic Regression, seperti pada tabel 12.

Tabel 12. Pengujian Cross Validation Logistic Regression

Fold Ke-	Accuracy	Precision	Recall	F1-Measure
1	89.62%	90.42%	89.62%	89.36%
2	86.92%	87.37%	86.92%	86.56%
3	85.77%	86.69%	85.77%	85.51%
4	87.31%	87.40%	87.31%	87.02%
5	88.85%	89.65%	88.85%	88.61%
6	82.69%	82.74%	82.69%	82.22%
7	89.62%	90.04%	89.62%	89.40%
8	87.69%	87.69%	87.69%	87.55%
9	89.58%	89.52%	89.58%	89.53%
10	89.19%	89.71%	89.19%	89.10%
Rerata	87.74%	88.28%	87.92%	87.58%

Berdasarkan Tabel 12, hasil pengujian K-Fold Cross Validation pada algoritma Logistic Regression menunjukkan bahwa performa tertinggi tercapai pada iterasi pertama dengan nilai akurasi 89.62%, presisi 90.42%, recall 89.62%, dan F-measure 89.36%. Detail hasil pengujian K-Fold Cross Validation pada algoritma Logistic Regression dapat dilihat pada Gambar 8 berikut.



Gambar 8. Grafik Cross Validation Logistic Regression

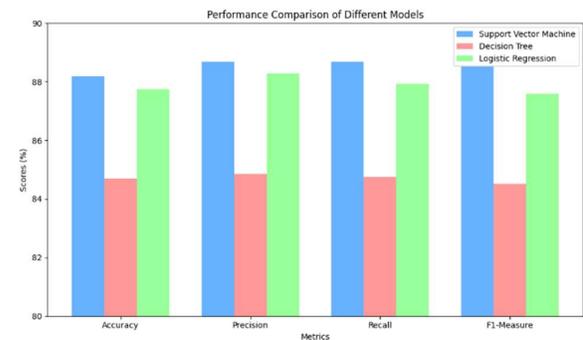
3.5. Perbandingan Pengujian Naïve Bayes dan K-Nearest Neighbor

Pada tahap ini, dilakukan perbandingan hasil antara dua algoritma yang digunakan dalam penelitian, yaitu algoritma Support Vector Machine, Decision Tree dan Logistic Regression. Perbandingan ini didasarkan pada hasil pengujian terakhir menggunakan Cross Validation. Hasil perbandingan kedua algoritma tersebut ditampilkan pada Tabel 13.

Tabel 13. Perbandingan Algoritma

Hasil Pengujian	Support Vector Machine	Decision Tree	Logistic Regression
Accuracy	88.18%	84.70%	87.74%
Precision	88.68%	84.86%	88.28%
Recall	88.68%	84.76%	87.92%
F1-Measure	88.51%	84.52%	87.58%

Berdasarkan Tabel 11, terlihat bahwa algoritma Support Vector Machine menunjukkan hasil pengujian yang lebih tinggi terhadap dokumen ulasan penelitian ini dibandingkan dengan algoritma Decision Tree dan Logistic Regression. Algoritma Support Vector Machine mencapai akurasi 88.51%, presisi 88.68%, recall 88.68%, dan F-measure 88.51%. Sementara itu, algoritma Decision Tree mencapai akurasi 84.70%, presisi 84.86%, recall 84.76%, dan F-measure 84.52% dan pada algoritma Logistic Regression mendapatkan accuracy 87.74%, precision 88.28%, recall 87.92%, dan F1-Measure 87.58%. Detail perbandingan hasil antara algoritma Support Vector Machine, Decision Tree, dan Logistic Regression dapat dilihat pada Gambar 7 berikut.



Gambar 9. Diagram Perbandingan Algoritma

4. PEMBAHASAN

Pada hasil penelitian diatas menunjukkan bahwa Support Vector Machine (SVM) secara konsisten memberikan performa yang lebih baik dalam klasifikasi sentimen dibandingkan Decision Tree, dengan SVM mencapai akurasi rata-rata 88.18% dan nilai tertinggi pada K-Fold Cross Validation mencapai 92.69%. Decision Tree, meskipun juga efektif, memiliki akurasi rata-rata 84.70%. Untuk melihat lebih detail klasifikasi sentimen pada aplikasi netflix, dapat dilihat pada berikut ini:

4.1. Persebaran dan Analisis Kata dalam Kelas Sentimen

Kelas sentimen dibagi menjadi dua kategori, yaitu positif dan negatif, pada ulasan pengguna aplikasi Netflix. Terdapat total 1.749 ulasan positif dan 1.138 ulasan negatif.

4.1.1. Sentimen Negative

Data sentimen negatif dari ulasan pengguna aplikasi Netflix berjumlah 1.138. Selanjutnya, dilakukan visualisasi distribusi kata dengan menggunakan Word Cloud. Hasil visualisasi ini ditampilkan pada Gambar 10.

Melakukan eksplorasi menggunakan teknik deep learning seperti Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), atau Transformer-based models seperti BERT atau GPT untuk klasifikasi sentimen. Pendekatan ini dapat menghasilkan hasil yang lebih baik dalam memahami nuansa bahasa dalam ulasan yang kompleks.

Analisis Multilingual: Mengembangkan model klasifikasi sentimen yang dapat menangani ulasan dalam berbagai bahasa, mengingat Netflix adalah platform global dengan pengguna dari berbagai negara dan latar belakang bahasa.

Penggunaan Data yang Lebih Luas: Memperluas dataset untuk mencakup lebih banyak ulasan pengguna dari berbagai platform, bukan hanya dari Google Play Store, tetapi juga dari platform sosial media lainnya seperti Twitter, Facebook, atau Reddit. Hal ini dapat memberikan gambaran yang lebih komprehensif tentang sentimen pengguna terhadap layanan Netflix.

Penggunaan Metrik Evaluasi Tambahan: Selain akurasi, precision, recall, dan F-measure, mengeksplorasi penggunaan metrik evaluasi seperti ROC-AUC atau metrik yang lebih sesuai untuk data tidak seimbang (imbalanced data) untuk mendapatkan pemahaman yang lebih baik tentang kinerja model klasifikasi.

DAFTAR PUSTAKA

- [1] O. Hondro, "Analisis Sentimen Pengguna Aplikasi E-Wallet Dana Melalui Postingan di Media Sosial Twitter Menggunakan Naïve Bayes," *KETIK J. Inform.*, 2023, [Online]. Available: <https://jurnal.faatuatua.com/index.php/KETIK/article/view/18>.
- [2] K. P. J. Sitompul, A. R. Pratama, and K. A. Baihaqi, "Komparasi Algoritma Naive Bayes, Support Vector Machine, dan Logistic Regression Pada Analisis Sentimen Pengguna Aplikasi Transportasi Online," *Kumpul. J. Ilmu Komput.*, vol. 10, no. 01, pp. 27–38, 2023.
- [3] A. Zamsuri, E. Asril, and M. Sadar, "Analisis Sentimen Opini Terhadap Vaksin Covid-19 Pada Media Sosial Twitter," vol. 5, no. 1, pp. 100–110, 2015.
- [4] M. I. Amal, E. S. Rahmasita, E. Suryaputra, and N. A. Rakhmawati, "Analisis Klasifikasi Sentimen Terhadap Isu Kebocoran Data Kartu Identitas Ponsel di Twitter," *J. Tek. Inform. dan Sist. Inf.*, vol. 8, no. 3, 2022, doi: [10.28932/jutisi.v8i3.5483](https://doi.org/10.28932/jutisi.v8i3.5483).
- [5] F. A. Larasati, D. E. Ratnawati, and ..., "Analisis Sentimen Ulasan Aplikasi Dana dengan Metode Random Forest," *J. Pengemb.*, ..., 2022, [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/11562>.
- [6] N. N. Sepriadi, "Analisis Sentimen Review Aplikasi My Pertamina Menggunakan Word Embedding Fasttext Dan Algoritma K-Nearest Neighbor," *Anal. Sentimen Rev. Apl.*, ..., 2023, [Online]. Available: <https://repository.uin-suska.ac.id/73902/>.
- [7] A. A. Achmad, K. Iin, and Y. Iska, "Analisis Klasifikasi Sentimen Berbasis Topik pada Ulasan Layanan Dana dan Sakuku dengan Convolutional Neural Network," *Inf. (Jurnal Inform. dan ...)*, 2023, [Online]. Available: <http://ojs.stmik-im.ac.id/index.php/INFORMASI/article/view/267>.
- [8] Oktavianus and M. Hondro, "Analisis Sentimen Pengguna Aplikasi E-Wallet Dana Melalui Postingan di Media Sosial Twitter Menggunakan Naïve Bayes," *J. Inform.*, vol. 01, no. 01, pp. 27–31, 2023.
- [9] I. N. Kusuma and I. Ali, "Analisis Sentimen Pada Pengguna Aplikasi Dana Menggunakan Algoritma Naive Bayes," *JATI (Jurnal Mhs. Tek. Inform.)*, 2024, [Online]. Available: <https://ejournal.itn.ac.id/index.php/jati/article/view/9041>.
- [10] W. E. Saputro, H. Yuana, and ..., "Analisis Sentimen Pengguna Dompot Digital Dana Pada Kolom Komentar Google Play Store Dengan Metode Klasifikasi Support ...," *JATI (Jurnal Mhs. ...)*, 2023, [Online]. Available: <https://ejournal.itn.ac.id/index.php/jati/article/view/6842>.
- [11] A. SAPUTRA, *Analisis Sentimen Aplikasi Investasi Reksa Dana (Bibit) Menggunakan Metode Support Vector Machine*. eprints.upnyk.ac.id, 2023.
- [12] M. et Al., *Introduction to Information Retrieval*. 2008.
- [13] F. F. Kiedrowsky and Andrianingsih, "Sentiment Analysis Marketplaces Digital menggunakan Machine Learning," *J. JTik (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 7, no. 3, pp. 493–499, 2023, doi: [10.35870/jtik.v7i3.1002](https://doi.org/10.35870/jtik.v7i3.1002).
- [14] Imamah and F. H. Rachman, "Twitter sentiment analysis of Covid-19 using term weighting TF-IDF and logistic regression," *Proceeding - 6th Inf. Technol. Int. Semin. ITIS 2020*, pp. 238–242, 2020, doi: [10.1109/ITIS50118.2020.9320958](https://doi.org/10.1109/ITIS50118.2020.9320958).
- [15] H. D. Wulandari and I. A. Rahman, "Cryptocurrency Price Volatility Analysis On Bitcoin And Altcoins Before And During The Covid-19 Pandemic In Indonesia," *J.*, ..., 2022, <https://doi.org/10.25077/TEKNOSI.v10i2.2024.110-117>

[Online]. Available:
<http://ejournal.aibpmjournals.com/index.php/JICP/article/view/2008>.

- [16] F. M. Delta Maharani, A. Lia Hananto, S. Shofia Hilabi, F. Nur Apriani, A. Hananto, and B. Huda, "Perbandingan Metode Klasifikasi Sentimen Analisis Penggunaan E-Wallet Menggunakan Algoritma Naive Bayes dan K-Nearest Neighbor," *Metik J.*, vol. 6, no. 2, pp. 97–103, 2022, doi: [10.47002/metik.v6i2.372](https://doi.org/10.47002/metik.v6i2.372).