



The 5th Conference on Innovation and Application of Science and Technology
(CIASTECH)

Website Ciastech 2022 : <https://ciastech.widyagama.ac.id>

Open Conference Systems : <https://ocs.widyagama.ac.id>

Proceeding homepage : <http://publishing-widyagama.ac.id/ejournal-v2/index.php/ciastech/index>

P-ISSN : 2622-1276

E-ISSN: 2622-1284

KINERJA SELEKSI FITUR N-GRAM PADA ANALISIS SENTIMEN ULASAN MOBILE GAME DI GOOGLE PLAYSTORE

Syahroni Wahyu Iriananda^{1*}, Renaldi Widi Budiawan², Aviv Yuniar Rahman³, Istiadi⁴

^{1,2,3,4} Program Studi S1 Teknik Informatika, Fakultas Teknik, Universitas Widyagama Malang

INFORMASI ARTIKEL

Data Artikel :

Naskah masuk, 8 September 2022

Direvisi, 30 September 2022

Diterima, 17 Oktober 2022

Email Korespondensi :

syahroni@widyagama.ac.id

ABSTRAK

Mobile Game menjadi industri yang berkembang pesat. Pengguna dan pengembang game menggunakan analisis sentimen untuk dapat mengetahui ulasan pemain game, sehingga dapat menetapkan arah pengembangan dan peningkatan game tersebut. Dalam penelitian ini, penulis mengklasifikasikan sentimen dengan menggunakan TF-IDF dan fitur seleksi N-Gram. Penulis bertujuan untuk menginvestigasi pengaruh N-Gram terhadap tingkat akurasi algoritma SVM. Maka dari itu, peneliti melakukan beberapa skenario pengujian penggunaan N-Gram. Hasil pengujian tersebut menunjukkan bahwa dengan menggunakan 3600 baris data, fitur seleksi Unigram, Bigram, dan Trigram (1,3), dan *cross-validation* $k=10$ dapat menghasilkan nilai akurasi yang optimum pada rasio data latih dan data uji 72:25 adalah 87,3 persen. Nilai ini dicapai dengan menggunakan kernel Radial Basis Function (RBF), dan parameter kompleksitas $C = 1$, Gamma (γ) = 1.

Kunci : N-Gram, Sentiment, Analysis, Playstore, Mobile Game

1. PENDAHULUAN

Permainan *mobile online game* telah menjadi suatu budaya pada sektor yang sedang berkembang. Pertumbuhan eksplosif *game* daring dewasa ini dapat didorong oleh perkembangan teknologi internet dan *smartphone*. Pada masa pandemi virus corona (COVID-19) lalu kebijakan pemerintah terkait jarak sosial (*social distancing*), membatasi aktivitas masyarakat dan bisnis. Mandat untuk tetap tinggal di rumah dan karantina menyebabkan partisipasi masyarakat dalam bermain game meningkat. Ini adalah salah satu cara untuk menghibur diri mereka yaitu bermain *Mobile game* secara online bersama teman mereka ataupun dengan orang asing [1]. Sebelum mereka mengunduh dan memasang *mobile game* pada perangkat, biasanya mereka memeriksa ulasan atau *review* dan peringkat game terlebih dahulu [2].

Dalam teks ulasan (*review*) pengguna di Google Playstore sering ditemukan kata-kata atau istilah yang tidak konsisten, tidak baku, menggunakan bahasa gaul atau *slang*, penggunaan istilah

rancu dan redundansi. Sebelum melakukan analisis sentimen diperlukan *preprocessing* teks [3], ini diperlukan untuk menangani ketidakseimbangan kelas yang dapat menyebabkan nilai akurasi dari proses klasifikasi sentimen menjadi tidak optimal [4].

Dalam proses klasifikasi sentiment sering menggunakan algoritma Support Vector Machine (SVM), algoritma ini cocok untuk digunakan dalam klasifikasi data teks karena sifat dasar teks yang cenderung mempunyai dimensi yang tinggi [5] dan mampu memperoleh tingkat akurasi tinggi dan mudah dilatih dibandingkan dengan algoritma pembelajaran mesin lainnya. Penggunaan memori lebih efisien karena keunggulan pemetaan kernel ke ruang fitur dimensi tinggi [6]. SVM juga dapat menangani berbagai macam kasus dan skenario dalam tugas *text mining* yang menggunakan Bahasa Indonesia [7]. Merujuk pada penelitian terdahulu, beberapa kajian dengan algoritma SVM yang ditulis oleh [8], [9], [10], dan [11] dengan nilai akurasi rata-rata diatas 81%, maka penulis bermaksud menggunakan metode ini sebagai algoritma pengklasifikasi sentiment.

Dalam ruang berdimensi tinggi ini dilakukan transformasi fitur (*feature selection*) untuk meningkatkan hasil klasifikasi yang lebih akurat dengan N-Gram yang diterapkan dengan melakukan modifikasi pemecahan berdasarkan kata [12]. Penerapan N-Gram pada ekstraksi TFIDF menjadikan algoritma perhitungan TFIDF menjadi lebih efektif [13]. Proses *preprocessing* dan kombinasi N-Gram mampu menghasilkan akurasi sebesar 96% pada tahap pengujian [14].

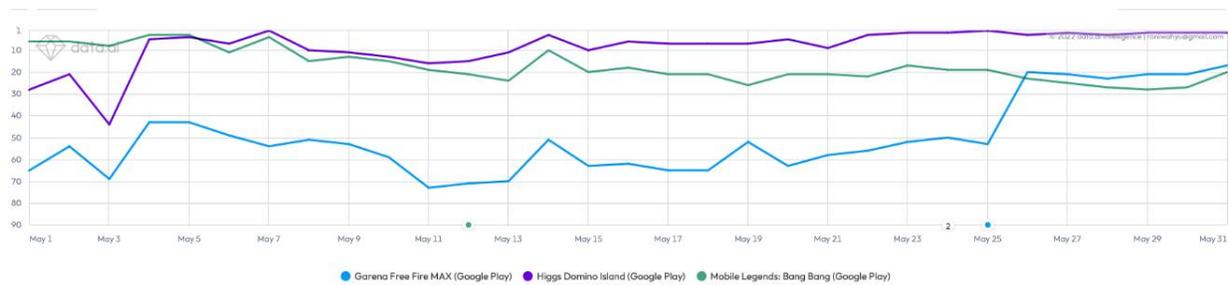
Penelitian yang menggunakan N-Gram dilakukan oleh [15], [16], [17], [18] dengan kinerja N-Gram pada beberapa penelitian terkait memiliki hasil yang memuaskan. Pada penelitian [15] kinerja Unigram (Uni) adalah 86%, nilai ini lebih baik daripada Bigram (Bi). Sementara itu pada penelitian [16], Unigram (Uni) menghasilkan akurasi sebesar 83,04%, indikator ini menunjukkan Unigram (Uni) lebih baik daripada Bigram (Bi) dan Trigram (Tri). Sedangkan pada penelitian [17] menggunakan SVM kernel RBF dengan Unigram & Bigram (UniBi) menghasilkan nilai akurasi yang lebih baik yaitu 93% daripada menggunakan Unigram (Uni) yaitu 90,43%. Jumlah data berpengaruh terhadap kinerja N-Gram ditunjukkan oleh penelitian [18] dengan menggunakan 900 baris data, nilai akurasi optimum dapat dicapai dengan menggunakan Unigram (Uni), namun dengan menggunakan 1200 baris data nilai akurasi lebih baik dicapai dengan menggunakan Trigram (Tri) [18].

Penelitian terdahulu yang menggunakan objek ulasan *game mobile* dilakukan oleh [4] menggunakan metode BERT sebagai model kecerdasan buatan untuk menganalisa sentiment game Genshin Impact menghasilkan nilai akurasi 74%, sementara [19] menggunakan metode Naïve Bayes untuk menganalisa sentimen game Mobile Legends: Bang Bang (MLBB) menghasilkan nilai akurasi 80%, [20] menganalisa game Harry Potter menghasilkan nilai akurasi 88%. Penelitian [21] menganalisa sentimen game MLBB, AoV dan Vainglory menghasilkan nilai akurasi 85,42% dengan menggunakan Naïve Bayes.

Dalam penelitian ini, penulis melakukan investigasi pengaruh N-Gram dan melakukan perbandingan kinerja beberapa model N-Gram pada studi kasus analisis sentimen ulasan (*review*) pengguna *mobile game* di Google Playstore menggunakan algoritma Support Vector Machine (SVM).

2. METODE PENELITIAN

Langkah pertama yang dilakukan oleh penulis adalah melakukan observasi terhadap *mobile game* dengan tujuan untuk mendapatkan data yang relevan dengan kajian. Observasi dilakukan data pada website Data.ai dengan kriteria observasi data adalah *mobile game* terpopuler dengan pendapatan paling tinggi dan dengan jumlah unduhan paling banyak di Google Play Store yang dapat diunduh secara gratis dan digunakan pada perangkat Android, dengan judul kategori Game.



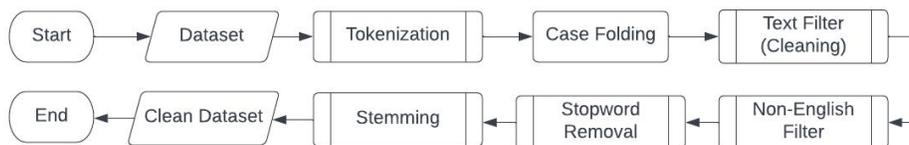
Gambar 1. Observasi Mobile Game populer di Google Playstore

Merujuk pada Gambar 1, dalam kurun waktu tertentu aplikasi game online Mobile Legend: Bang Bang (MLBB) memiliki tren positif dengan peringkat 10 Besar untuk aplikasi kategori Game, namun sejak awal bulan game ini sudah tergeser oleh game Higgs Domino Island (HDI). Posisi berikutnya adalah game HDI yang di awal bulan berada di peringkat 30 Besar. Dalam kurun waktu 2 hari meningkat tajam masuk ke peringkat 5 Besar sampai akhir bulan Mei. Yang perlu menjadi perhatian berikutnya adalah game Garena Free Fire Max (GFFM) dimana peringkatnya masuk di 70 Besar di awal bulan, kemudian mengalami peningkatan yang cukup signifikan pada akhir Bulan dan masuk di peringkat 25 Besar. Dari gambar tersebut memberikan indikator bahwa peringkat dari setiap aplikasi game online selalu fluktuatif.

Langkah kedua adalah mengumpulkan data ulasan dari Google Playstore. Dalam kajian ini penulis mengambil data *review* dari tiga game terpopuler di google playstore yaitu MLBB, GFFM, dan HDI. Data ketiga *review mobile game* tersebut didapatkan dan dikumpulkan melalui situs layanan *web scraping* yaitu *apify.com*. Peneliti menggunakan total dataset awal ini adalah sebanyak 7200 baris data yang terdiri dari 2400 baris data untuk masing-masing game. Dengan proporsi label "Positif" dan "Negatif" masing-masing sebanyak 1200 baris data. Hasil dataset tersebut kemudian digunakan dalam proses selanjutnya yaitu *Preprocessing teks*.

a. Preprocessing Teks

Dalam teks ulasan pengguna (*review*) di Google Playstore sering ditemukan pada kalimat atau kata-kata atau istilah yang tidak konsisten, tidak baku, penggunaan bahasa gaul atau *slang*, penggunaan istilah yang tidak jelas atau rancu dan redundansi, sehingga sulit untuk melakukan data *preprocessing* sebelum melakukan analisis sentimen [3]. Tahap *preprocessing* memerlukan banyak waktu dan sumber daya, selain generalisasi data, penapisan, normalisasi, konstruksi, dan transformasi data yang dilakukan selama fase *preprocessing*. Hal ini juga diperlukan untuk menangani ketidakseimbangan, karena kelas yang tidak seimbang akan menyebabkan nilai akurasi dari proses klasifikasi yang tidak tepat [4]. Gambaran global *preprocessing* teks dapat dilihat pada Gambar 2 berikut

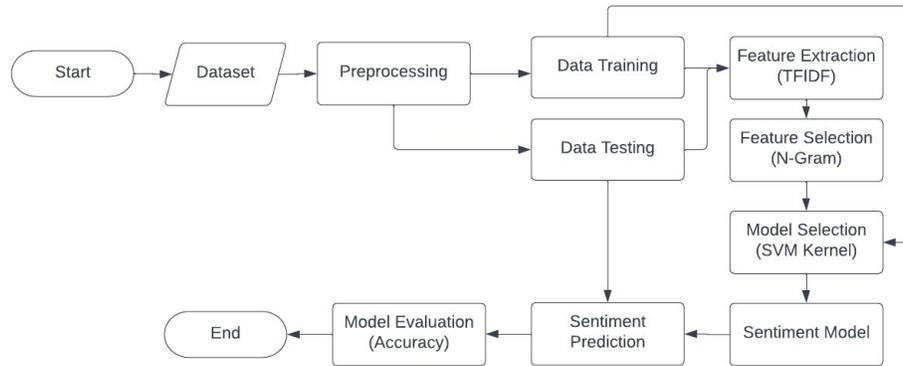


Gambar 2. Proses *Preprocessing* Teks

b. Model Analisis Sentimen

Proses analisis sentimen diawali dengan (input) dataset ulasan pengguna *mobile game* populer pada Google Play Store. Kemudian proses *preprocessing* teks yang hasilnya dibagi menjadi dua bagian, yaitu data *Training*, dan data *Testing*. Proses Selanjutnya adalah ekstraksi fitur dengan pembobotan kata menggunakan *TFIDF*, kemudian dipilih fitur terbaik dengan N-Gram. Dalam

proses ini juga dapat aplikasikan untuk pencarian jangkauan N-Gram (*n-gram range*) dan batasan fitur maksimal (*max feature*) untuk membatasi jumlah fitur dalam ruang dimensi tinggi. Selanjutnya digunakan pada proses seleksi model dengan memanfaatkan data *Training*. Secara global model analisis sentimen yang diusulkan dapat dilihat pada Gambar 3.



Gambar 3. Model Analisis Sentimen

Proses *Model Selection* menggunakan optimasi *hyperparameter* dengan metode *Grid Search (GS)* guna mendapatkan *hyperparameter* SVM yang optimal. Proses *Sentiment Prediction* memanfaatkan data *Testing* yang telah ditentukan dan dibagi dari proses sebelumnya. Prediksi sentimen ini menggunakan model SVM yang optimal pada proses sebelumnya, sehingga didapatkan nilai akurasi yang optimal pada proses evaluasi. Evaluasi dalam kajian ini menggunakan Confusion Matrix

c. Evaluasi Confusion Matrix

Confusion matrix adalah tabel yang digunakan untuk mendefinisikan kinerja dari algoritma klasifikasi, memvisualisasikan dan merangkum kinerja algoritma klasifikasi [22]. Confusion Matrix ditunjukkan pada Tabel 2 berikut.

Tabel 2. Confusion Matrix

Actual Value	Predicted value	
	Positif	Negatif
Positif	TP	FN
Negatif	FP	TN

Dalam pembelajaran mesin dan ilmu data, pengukuran atau metrik paling populer yang digunakan untuk mengevaluasi model klasifikasi adalah Confusion Matrix yang terdiri dari empat karakteristik dasar (angka) yang digunakan untuk menentukan metrik pengukuran algoritma pengklasifikasi yang antara lain: 1) *True Positive (TP)*: Mewakili jumlah label yang telah diklasifikasikan dengan benar memiliki sentimen Positif, artinya mereka memiliki label tersebut. 2) *True Negatif (TN)* mewakili jumlah label Negatif yang diklasifikasikan dengan benar. 3) *False Positive (FP)* mewakili jumlah label Positif yang salah diklasifikasikan Negatif namun sebenarnya mereka adalah Positif. 4) *False Negative (FN)* mewakili jumlah label Negatif diklasifikasikan sebagai Positif namun tetapi sebenarnya mereka adalah Negatif. Metrik kinerja suatu algoritme adalah akurasi, precision, recall, dan skor F1, yang dihitung berdasarkan TP, TN, FP, dan FN yang disebutkan di atas [22]. Formula metrik kinerja tersebut disajikan pada rumus Akurasi (1), Precision (2), Recall (3), dan skor F1 (4) berikut:

$$Accuracy = \frac{(TP+TN)}{(TP+FP+FN+TN)} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

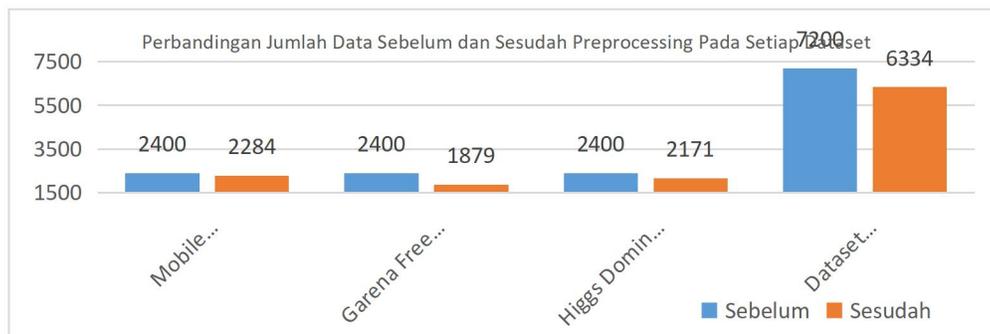
$$F1 \text{ Score} = 2 \times \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad (4)$$

Akurasi (*accuracy*) memberikan proporsi jumlah total prediksi yang benar. Precision atau nilai prediksi positif, adalah fraksi nilai positif dari total kejadian positif yang diprediksi. Dengan kata lain, presisi adalah proporsi nilai positif yang diidentifikasi dengan benar. Recall disebut juga *Sensitivity*, atau TP rate (TPR) adalah bagian dari nilai positif dari total kasus positif aktual yaitu, proporsi kasus positif aktual yang diidentifikasi dengan benar. Skor F1, skor F, atau *F-measure* adalah rata-rata harmonik presisi dan recall yang memberikan arti penting bagi kedua faktor [23].

3. HASIL DAN PEMBAHASAN

a. Statistik Hasil Proses Preprocessing Teks

Pada proses preprocessing teks dilakukan berbagai macam filter teks sehingga berpengaruh terhadap jumlah baris data. Perbandingan jumlah baris dataset dipresentasikan pada Gambar 4.



Gambar 4. Statistik Perbandingan Jumlah Dataset Sebelum dan Sesudah Preprocessing

Hal ini terjadi karena hasil filter teks tersebut berpotensi menghapus seluruh kata sesuai filter yang diaplikasikan sehingga menjadi data kosong (*missing value*), jumlah dataset sebelum dilakukan proses preprocessing adalah 7200 baris data berkurang menjadi 6334 baris data. Pada Gambar 4 tersebut menunjukkan perbandingan antara jumlah baris data dalam dataset sebelum dilakukan proses *preprocessing* dan sesudah dilakukan proses *preprocessing*, dimana dataset Garena Free Fire Max (GFFM) memiliki jumlah dataset yang paling minimum yaitu 1879, sedangkan jumlah baris data maksimum pada dataset MLBB yaitu 2284 baris data. Guna menghindari ketidak-seimbangan (*imbalanced*) pada setiap dataset, dalam penelitian ini penulis hanya menggunakan 3600 baris dataset dengan proporsi 1800 baris data dengan label “positif” dan “negatif” yang sama. Pada Tabel 2 ditunjukkan jumlah label pada setiap dataset setelah dilakukan proses *preprocessing* teks.

Tabel 2. Statistik Jumlah Label Positif dan Negatif dalam Dataset setelah proses *preprocessing*

Dataset	Positif	Negatif
Mobile Legends : Bang Bang	1165	1119
Garena Free Fire Max	1009	870
Higgs Domino Island	1087	1084
Dataset Gabungan	3261	3073

Jumlah ini 1800 baris data setiap label tersebut mewakili 600 baris data dari setiap datasetnya. 600 baris data ditentukan untuk mengakomodasi jumlah label “Negatif” yang minimum yaitu pada dataset GFFM yaitu 870 data agar bisa diambil sampel secara merata.

b. Rasio Data Latih dan Data Uji

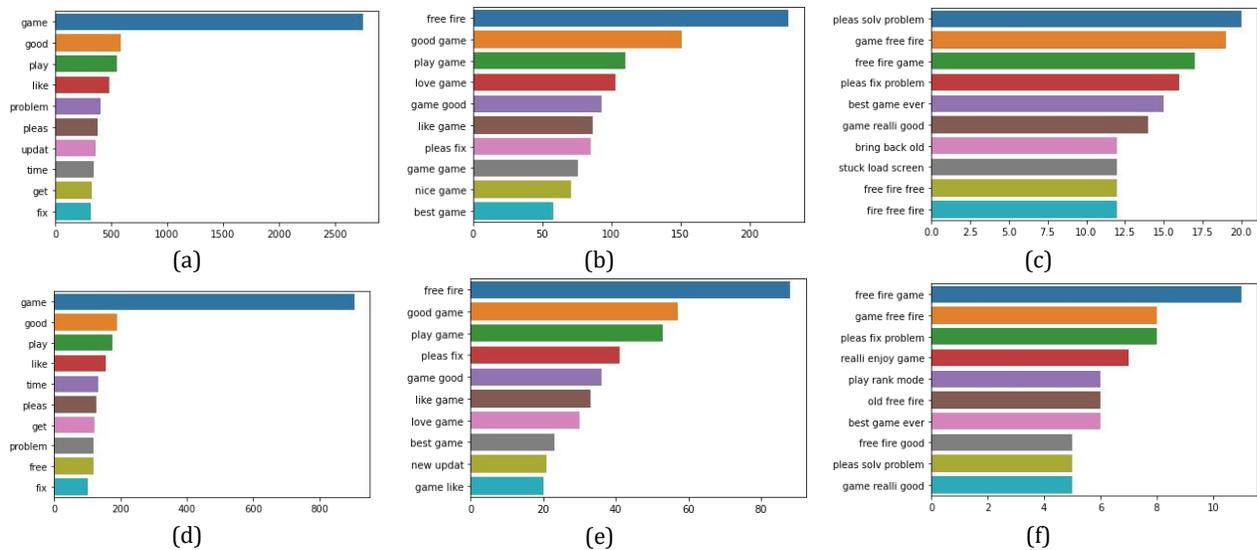
Pada pengujian tahap ini menggunakan 3600 baris data dari dataset gabungan (Allgame) dengan rasio perbandingan data latih dan data uji disajikan pada Tabel 1 berikut:

Tabel 1. Jumlah Rasio Perbandingan Label Positif dan Negatif

Label	Jumlah	Rasio 75:25
Positif	1800	1350:450
Negatif	1800	1350:450
Total	3600	2700:900

Label Positif dan Negatif memiliki jumlah yang sama yaitu total 1800 baris data pada setiap rasio perbandingan data latih dan data uji. Seperti yang disajikan pada tabel tersebut diatas, pada rasio 75:25 memiliki 1350 data latih dan 450 data uji.

c. Perbandingan Jumlah N-Gram



Gambar 5. Perbandingan N-Gram pada dataset Training dan Testing

Jumlah frekuensi N-Gram terbanyak dalam penelitian ini menunjukkan jumlah frekuensi kata yang muncul dalam dataset training dan data testing. Frekuensi ini digolongkan menjadi tiga bagian yaitu Unigram (Uni), Bigram (Bi) dan Trigram (Tri), kemudian dipilih 10 rangking frasa yang memiliki jumlah frekuensi terbanyak pada setiap model N-Gramnya. Frekuensi N-gram disajikan pada Gambar 5. Frekuensi N-gram pada dataset training berada pada bagian atas gambar (5a, 5b, 5c) , dan frekuensi N-Gram pada dataset testing berada pada bagian bawah gambar (5d, 5e, 5f). Frekuensi Unigram terdapat pada gambar (5a, 5d), frekuensi Bigram pada gambar (5b, 5e), dan frekuensi Trigram pada gambar (5c, 5f). Berdasarkan gambar tersebut, dapat diketahui bahwa kata “game” memiliki frekuensi terbanyak pada model Unigram (Uni), diikuti oleh kata “good” dan “play” secara berurutan, sedangkan pada model Bigram (Bi) jumlah frekuensi terbanyak adalah frasa “free fire”, diikuti oleh frasa “good game” dan “play game”. Hasil yang berbeda dapat dilihat pada model Trigram (Tri) frasa “pleas solv problem” muncul dengan frekuensi paling banyak pada dataset training, sedangkan “free fire game” merupakan frasa terbanyak kemunculannya pada dataset testing.

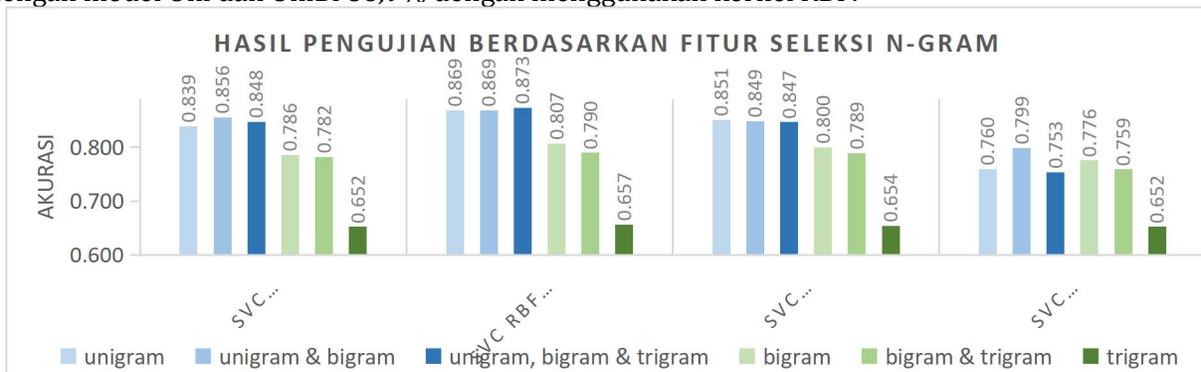
d. Kinerja N-Gram

Pengujian kinerja N-Gram pada klasifikasi analisis sentimen menggunakan metode SVM dilakukan dengan menggunakan gabungan seluruh dataset MLBB, GFFM, dan HDI kemudian disebut sebagai dataset gabungan (Allgame). Peramalan klasifikasi (Estimator SVC) kernel SVM menggunakan empat kernel yaitu Linear, Polynomial, RBF dan Sigmoid. Pada pengujian ini penulis menggunakan 3600 baris data dari dataset gabungan (Allgame) dengan rasio perbandingan data latih dan data uji adalah 75:25 yaitu 75% data latih dan 25% data uji, dengan *Feature Selection (FS)* menggunakan Hanya Unigram (Uni), model N-Gram Unigram dan Bigram (UniBi), model N-Gram Unigram, Bigram dan Trigram (UniBiTri), Hanya Bigram (Bi), Bigram dan Trigram (BiTri), Hanya Trigram (Tri) dengan 3000 fitur maksimal. Dalam proses klasifikasi sentiment menggunakan cross-validation KFold = 10. hasil pencarian *hyperparameter* dengan Grid Search (GS) memberikan rekomendasi nilai Gamma (γ) adalah Gamma = 1 dan nilai C (complexity) adalah C = 1 dengan nilai. Hasil pengujian seleksi fitur N-Gram terdapat pada Tabel 3 berikut:

Tabel 3. Kinerja Model N-Gram

Kernel SVM	N-Gram					
	Uni	UniBi	UniBiTri	Bi	BiTri	Tri
Linear	0.839	0.856	0.848	0.786	0.782	0.652
RBF	0.869	0.869	0.873	0.807	0.790	0.657
Polynomial (degree 3)	0.851	0.849	0.847	0.800	0.789	0.654
Sigmoid	0.760	0.799	0.753	0.776	0.759	0.652

Dalam Tabel 3 disajikan nilai akurasi model N-Gram Hanya Unigram (Uni), model N-Gram Unigram dan Bigram (UniBi), model N-Gram Unigram, Bigram dan Trigram (UniBiTri), Hanya Bigram (Bi), Bigram dan Trigram (BiTri), Hanya Trigram (Tri). Pada Tabel tersebut dapat dilihat bahwa nilai akurasi maksimum didapatkan dengan model N-Gram UniBiTri adalah 87,3% diikuti dengan model Uni dan UniBi 86,9% dengan menggunakan kernel RBF.



Gambar 6. Perbandingan Kinerja Model N-Gram

Dalam pengujian ini dapat dilihat bahwa kernel RBF selalu mendapatkan nilai akurasi tertinggi dibandingkan dengan kernel yang lainnya. Nilai akurasi minimum didapatkan dari seleksi fitur hanya Trigram (Tri) yaitu 65,2% dengan kernel Linear dan Sigmoid. Dari grafik pada Gambar 6 tersebut fitur seleksi Bigram (Bi), Bigram dan Trigram (BiTri) dan Trigram (Tri) rata-rata memiliki angka akurasi yang lebih rendah daripada seleksi fitur yang melibatkan Unigram didalamnya. Namun hal yang berbeda pada kernel Sigmoid, dimana nilai akurasi fitur seleksi bigram (Bi) lebih baik daripada unigram (Uni). Berdasarkan hasil pengujian fitur seleksi N-Gram ini maka penulis mengambil kesimpulan bahwa fitur seleksi N-Gram memiliki pengaruh pada kinerja klasifikasi sentimen menggunakan algoritma SVM.

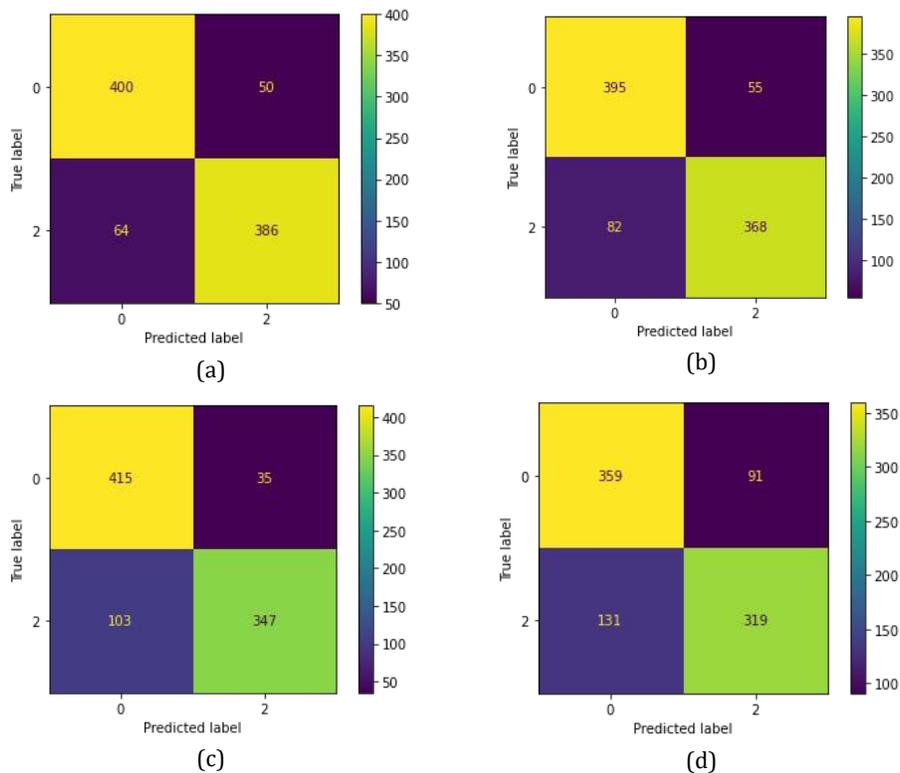
e. Kinerja Klasifikasi SVM

Pada penelitian ini nilai akurasi terbaik berdasarkan Tabel 3 adalah model N-Gram UniBiTri dengan menggunakan kernel RBF yaitu 87,3%. Kinerja Kernel RBF tersebut ditunjukkan pada laporan klasifikasi pada Gambar 7 dan Confusion Matrix pada Gambar 8 berikut:

	precision	recall	f1-score	support
0	0.89	0.86	0.88	464
2	0.86	0.89	0.87	436
accuracy			0.87	900
macro avg	0.87	0.87	0.87	900
weighted avg	0.87	0.87	0.87	900

Gambar 7. Classification Report Kernel RBF

Berdasarkan laporan klasifikasi pada Gambar 7 tersebut dapat ditunjukkan bahwa nilai precision atau proporsi nilai positif yang diidentifikasi dengan benar paling optimum adalah pada kelas 0 atau “Negatif” yaitu 89%, untuk kelas 2 atau “Positif” adalah 86%. Sedangkan nilai Recall, proporsi kasus positif aktual yang diidentifikasi dengan benar paling optimum adalah pada kelas Positif yaitu 89%. Dan nilai rata-rata harmonik presisi dan recall yang memberikan arti penting bagi kedua faktor atau skor F1 optimum adalah 88% pada kelas Negatif. Dengan nilai support pada kelas Negatif adalah 464 dan kelas Positif adalah 436. Akurasi dan nilai makro pada laporan Kernel RBF tersebut menunjukkan 87% dengan nilai support 900.



Gambar 8. Evaluasi Confusion Matrix Kernel SVM

Performansi algoritma pengklasifikasi SVM dalam penelitian ini diukur dengan confusion matrix ditunjukkan pada Gambar 8. Confusion matrix algoritma SVM dengan Kernel RBF

ditunjukkan pada Gambar 8a, kernel Linear pada Gambar 8b, kernel Polynomial dengan derajat (degree) = 3 pada Gambar 8c, dan Sigmoid pada Gambar 8d.

4. KESIMPULAN

Berdasarkan pengujian yang telah dilakukan dapat disimpulkan bahwa Algoritma SVM dapat diterapkan untuk tugas analisis sentiment pada teks ulasan pengguna mobile game online populer di Google Play Store. Pemilihan *hyperparameter* menggunakan algoritma Grid Search dan Cross-Validation dengan $cv = 10$ menghasilkan *best parameter C (complexity)* dan Γ pada algoritma SVM yang tepat dapat menghasilkan nilai akurasi yang optimal. Hasil rekomendasi GS tersebut memberikan pemilihan kernel yang tepat yaitu kernel Radial Basis Function (RBF) menggunakan 3600 baris data dari dataset gabungan (allgame), dengan nilai parameter C (complexity) adalah 1 (satu) dan Γ adalah 1 (satu). Kinerja model N-Gram dapat meningkatkan nilai akurasi algoritma klasifikasi. Pada penelitian ini menggunakan feature selection UniBiTri dengan fitur maksimal adalah 3000 dapat menghasilkan nilai akurasi optimum yaitu 87,3% menggunakan kernel RBF, dan 3600 baris data dari dataset gabungan (allgame). Nilai akurasi tertinggi didapatkan dengan model N-Gram UniBiTri adalah 87,3% diikuti dengan model Uni dan UniBi 86,9% dengan menggunakan kernel RBF. Sedangkan nilai akurasi terendah didapatkan dari seleksi fitur hanya Trigram (Tri) yaitu 65,2% dengan kernel Linear dan Sigmoid.

Pada penelitian selanjutnya disarankan untuk menggunakan jumlah data yang lebih banyak dan kualitas dataset yang lebih baik, misalnya lebih dari 5.000 baris data agar dapat meningkatkan performa algoritma. Penggunaan *preprocessing* yang tepat juga dapat meningkatkan akurasi karena dengan dataset yang berkualitas dan lebih banyak mengandung kata baku dapat meningkatkan kinerja klasifikasi. Pendekatan analisis sentimen berbasis leksikon sangat disarankan agar dapat memberikan hasil yang optimal misalnya dengan menggunakan *word embedding* Word2Vec, VADER dan TextBlob. Dalam penelitian ini belum menerapkan, *Pos Tagging*, dan *Word Lemmatizer* sehingga pada penelitian selanjutnya dapat mengganti fungsi Stemming dengan Pos Tagging dan Word Lemmatizer agar bisa mendapatkan hasil akurasi yang lebih baik.

5. UCAPAN TERIMA KASIH

Terima kasih kepada Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) Universitas Widyagama Malang yang telah mendukung penelitian ini melalui program Hibah Penelitian Internal (PERINTIS)

6. REFERENSI

- [1] D. L. King, P. H. Delfabbro, J. Billieux, and M. N. Potenza, "Problematic online gaming and the COVID-19 pandemic," *Journal of Behavioral Addictions*, vol. 9, no. 2, pp. 184–186, Jun. 2020, doi: 10.1556/2006.2020.00016.
- [2] S. Chakraborty, I. Mobin, A. Roy, and M. H. Khan, "Rating Generation of Video Games using Sentiment Analysis and Contextual Polarity from Microblog," in *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, Dec. 2018, pp. 157–161. doi: 10.1109/CTEMS.2018.8769149.
- [3] S. Pradha, M. N. Halgamuge, and N. Tran Quoc Vinh, "Effective Text Data Preprocessing Technique for Sentiment Analysis in Social Media Data," in *2019 11th International Conference on Knowledge and Systems Engineering (KSE)*, Oct. 2019, pp. 1–8. doi: 10.1109/KSE.2019.8919368.
- [4] R. Kusnadi, Y. Yusuf, A. Andriantony, R. Ardian Yaputra, and M. Caintan, "ANALISIS SENTIMEN TERHADAP GAME GENSHIN IMPACT MENGGUNAKAN BERT," *rabit*, vol. 6, no. 2, pp. 122–129, Jul. 2021, doi: 10.36341/rabit.v6i2.1765.

- [5] R. A. Nandini, Y. A. Sari, and P. P. Adikara, "Analisis Sentimen Impor Beras 2018 Pada Twitter Menggunakan Metode Support Vector Machine dan Pembobotan Jumlah Retweet," p. 11, 2019.
- [6] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowledge-Based Systems*, vol. 226, p. 107134, Aug. 2021, doi: 10.1016/j.knosys.2021.107134.
- [7] A. P. Thenata, "Text Mining Literature Review on Indonesian Social Media," *JEPIN*, vol. 7, no. 2, p. 226, Aug. 2021, doi: 10.26418/jp.v7i2.47975.
- [8] S. Ranjan and S. Mishra, "Comparative Sentiment Analysis of App Reviews," in *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, Kharagpur, India, Jul. 2020, pp. 1–7. doi: 10.1109/ICCCNT49239.2020.9225348.
- [9] M. I. Ahmadi, F. Apriani, M. Kurniasari, S. Handayani, and D. Gustian, "SENTIMENT ANALYSIS ONLINE SHOP ON THE PLAY STORE USING METHOD SUPPORT VECTOR MACHINE (SVM)," p. 8, 2020.
- [10] H. Samy, A. Helmy, and N. Ramadan, "Aspect-based Sentiment Analysis of Mobile Apps Reviews using Class Association Rules and LDA," in *2021 Tenth International Conference on Intelligent Computing and Information Systems (ICICIS)*, Dec. 2021, pp. 183–189. doi: 10.1109/ICICIS52592.2021.9694242.
- [11] S. W. Iriananda, R. P. Putra, and K. S. Nugroho, "ANALISIS SENTIMEN DAN ANALISIS DATA EKSPLORATIF ULASAN APLIKASI MARKETPLACE GOOGLE PLAYSTORE," Malang, Indonesia, 2021, p. 10. [Online]. Available: <http://publishing-widyagama.ac.id/ejournal-v2/index.php/ciastech/article/download/3343/1800>
- [12] N. Arifin, U. Enri, and N. Sulistiyowati, "Penerapan Algoritma Support Vector Machine (SVM) dengan TF-IDF N-Gram untuk Text Classification," *STRING*, vol. 6, no. 2, p. 129, Dec. 2021, doi: 10.30998/string.v6i2.10133.
- [13] S. Mardianti, M. Z. Nafan, and I. Hidayatulloh, "EKSTRAKSI TF-IDF N-GRAM DARI KOMENTAR PELANGGAN PRODUK SMARTPHONE PADA WEBSITE E-COMMERCE," p. 7, 2018.
- [14] I. Pujadayanti, M. A. Fauzi, and Y. A. Sari, "Prediksi Rating Otomatis pada Ulasan Produk Kecantikan dengan Metode Naïve Bayes dan N-gram," p. 7, 2018.
- [15] R. H. Faturrahman, W. Astuti, and M. D. Purbolaksono, "Klasifikasi Sentimen Ulasan Film Menggunakan Support Vector Machine, Information Gain, dan N-Grams," p. 6, 2022.
- [16] S. A. Anjani and A. Fauzan, "Implementasi n-Gram dalam Analisis Sentimen Masyarakat DIY Terhadap PSBB Jawa-Bali Jilid II Menggunakan," vol. 21, no. 2, p. 11, 2021.
- [17] M. Oktafani and P. T. Prasetyaningrum, "IMPLEMENTASI SUPPORT VECTOR MACHINE UNTUK ANALISIS SENTIMEN KOMENTAR APLIKASI TANDA TANGAN DIGITAL," *Jurnal Sistem Informasi dan Bisnis Cerdas*, vol. 15, no. 1, pp. 10–19, Mar. 2022, doi: 10.33005/sibc.v15i1.2697.
- [18] F. Fitriyani and T. Arifin, "PENERAPAN WORD N-GRAM UNTUK SENTIMENT ANALYSIS REVIEW MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (STUDI KASUS: APLIKASI SAMBARA)," *SISTEMASI*, vol. 9, no. 3, p. 610, Sep. 2020, doi: 10.32520/stmsi.v9i3.954.
- [19] D. Sinaga and C. Jatmoko, "ANALISIS SENTIMEN UNTUK MENGETAHUI KESAN PLAYER GAME MOBILE LEGENDS MENGGUNAKAN NAÏVE BAYES CLASSIFIER," p. 8, 2020.
- [20] P. A. A. Rahman, E. Ermatita, and H. N. Irmanda, "ANALISIS SENTIMEN ULASAN GAME HARRY POTTER: HOGWARTS MYSTERY PADA SITUS GOOGLE PLAY MENGGUNAKAN NAÏVE BAYES CLASSIFIER," *SEINASI-KESI*, vol. 4, no. 1, Art. no. 1, Dec. 2021.
- [21] 14611229 Virga Inayatullah, "ANALISIS SENTIMEN GAME MULTIPLAYER ONLINE BATTLE ARENA (MOBA) MENGGUNAKAN ALGORITMA NAÏVE BAYES (Studi Kasus: Game Mobile Legend, AoV, Vainglory)," Sep. 2018, Accessed: Aug. 22, 2022. [Online]. Available: <https://dspace.uui.ac.id/handle/123456789/11814>