

ANALISIS SENTIMEN FILM AGAK LAEN DENGAN KECERDASAN BUATAN: TEXT MINING METODE NAÏVE BAYES CLASSIFIER

Adinda Cahya Kamilla, Natasya Priyani, Ressa Priskila, Viktor Handrianus Pranatawijaya

Teknik Informatika, Universitas Palangka Raya

Jalan Yos Sudarso, Palangka, Kec. Jekan Raya, Kota Palangka Raya, Kalimantan Tengah 74874

adindack@mhs.eng.upr.ac.id

ABSTRAK

Analisis sentimen terhadap ulasan film di media sosial Twitter selalu muncul sebagai subjek yang menarik untuk diselidiki dalam penelitian. Studi ini bertujuan untuk melakukan analisis sentimen terhadap ulasan film "Agak Laen" di media sosial Twitter menggunakan metode klasifikasi Naïve Bayes Classifier. Data diambil melalui *web scraping* dan diproses melalui tahap *preprocessing*, termasuk pembersihan, *case folding*, normalisasi, penghapusan *stopword*, tokenisasi dan *stemming*. Setelah dilatih dan dievaluasi menggunakan KFold cross-validator dengan $k=5$, model menunjukkan akurasi sebesar 78%. Evaluasi menunjukkan bahwa model memiliki *precision* sebesar 85.51%, *recall* sebesar 85.57%, dan *f1-score* sebesar 82.71%. Visualisasi grafik distribusi sentimen dan *word cloud* juga dilakukan untuk memberikan gambaran yang lebih jelas tentang data. Hasil penelitian ini menunjukkan bahwa metode klasifikasi Naïve Bayes Classifier memiliki potensi yang baik dalam menganalisis sentimen ulasan film di Twitter.

Kata Kunci : Film, Naïve Bayes Classifier, Analisis Sentimen, Text Mining

1. PENDAHULUAN

Topik-topik yang berkaitan dengan film merupakan materi yang menarik untuk dibahas di media[1]. Dengan pertumbuhan industri kreatif, produksi film menjadi lebih cepat, meningkatkan minat di bidang perfilman di negara ini. Salah satu faktor yang mendorong peningkatan jumlah penggemar film adalah kemudahan menikmati film. Film ini memiliki banyak aspek keindahan yang membuatnya disukai. Seperti narasi cerita, teknik fotografi, dan efek visual.

Suatu ulasan atau *review* dapat menjelaskan penilaian terhadap elemen-elemen tersebut. Ulasan dapat mencakup *review* film, yang dapat menghasilkan perasaan positif atau negatif. Selain itu, penilaian juga memiliki nilai angka dengan skala tertentu sebagai hasil dari *review* tersebut dan dapat membantu pembaca memilih film mana yang akan mereka tonton. Analisis sentimen berfungsi untuk menetapkan perasaan dan rating suatu *review* [2].

Semua orang harus membaca *review* film untuk mendapatkan informasi tentang film sehingga mereka dapat memahami isi film. Sebuah *review* film dapat memberikan informasi tentang jalan cerita, aktor, konflik, dan kelebihan dan kekurangan film. Informasi ini kemudian digunakan untuk menentukan kualitas film, dengan demikian para penyuka film memiliki pengetahuan sejauh mana film tersebut berhak ditonton [3].

Hingga saat ini, jumlah pengguna media sosial, terutama twitter, terus meningkat. Dengan mencari *hashtag* yang sedang trending, setiap orang dapat mengakses informasi dengan cepat. Peneliti *text mining* sangat tertarik untuk menggunakan analisis data media sosial twitter karena memungkinkan programmer untuk mendapatkan data secara bebas. Twitter juga unggul karena mudah digunakan dan

disesuaikan, alat penelitian yang kuat, mesin pencari yang *real-time*, dan analisis data yang cepat. Analisis sentimen adalah salah satu jenis analisis data yang paling umum [4].

Review penonton film "Agak Laen" di media sosial Twitter biasanya positif, negatif, atau netral. Hal ini pasti akan mempengaruhi ekspektasi penonton. Dibutuhkan sistem yang dapat menyortir dan mengategorikan ulasan secara cepat dan otomatis.

Text mining merupakan metode analisis data yang digunakan untuk mengekstraksi informasi penting dari teks untuk mengidentifikasi dan menganalisis pola, tren, dan topik yang relevan dengan topik tertentu. Ini dapat dilakukan dengan menggunakan analisis sentimen untuk mengidentifikasi sentimen yang terdapat pada pernyataan. Dalam konteks ini, pengembangan dan penerapan analisis sentimen sangat didorong oleh berbagai manfaat dan dampak positif yang ditawarkannya. [5].

Untuk analisis sentimen, menggunakan metode klasifikasi Naive Bayes. Metode ini cukup efektif dan mudah digunakan, sangat efisien, dan sangat akurat[6]. Analisis sentimen dari ulasan film diharapkan akan memberikan kemudahan bagi para pecinta film dalam menilai kualitas sebuah film.

2. TINJAUAN PUSTAKA

2.1. Text Mining

Text Mining merupakan cabang dari data mining yang memiliki tujuan serupa, yaitu mengekstrak data untuk menemukan pola atau hubungan unik yang mencerminkan isi atau karakteristik khusus dari sebuah dokumen teks untuk beberapa tujuan khusus[7]. Analisis sentimen, pembubaran dokumen, klasifikasi dokumen, ekstraksi informasi, pencarian

informasi, dan pemrosesan *web* adalah beberapa permasalahan yang dapat diatasi oleh *text mining*.

Text Mining adalah proses ekstraksi informasi dari data teks dan sumber data lainnya, yang umumnya berasal dari dokumen. Yang memungkinkan analisis hubungan antar dokumen, tujuan *text mining* adalah untuk menemukan kata-kata yang mewakili isi dokumen [8].

2.2. Analisis Sentimen

Analisis sentimen adalah bidang ilmu yang menganalisis pendapat, perilaku, dan emosi seseorang terhadap suatu entitas, seperti produk, layanan, organisasi, individu, isu, peristiwa, dan topik. Analisis sentimen berfokus pada pendapat yang mengekspresikan sentimen positif atau negatif. Karena manfaatnya yang besar, penelitian dan aplikasi berbasis analisis sentimen telah berkembang pesat [9].

2.3. Naïve Bayes Classifier

Teori tentang probabilitas penentuan kelompok dikenal sebagai Naïve Bayes Classifier. Algoritma klasifikasi ini memiliki kemampuan untuk mengolah sejumlah besar data dengan tingkat akurasi yang sangat tinggi. Beberapa keunggulan Naïve Bayes Classifier termasuk proses komputasi yang cepat, mudah digunakan dengan struktur yang sederhana, dan efektif. Saat digunakan untuk dataset berukuran besar, klasifikasi Naïve Bayes menunjukkan kecepatan dan akurasi yang tinggi. Karena strukturnya yang sederhana, klasifikasi Naïve Bayes sering digunakan dalam data *mining* [10].

Sebuah penelitian sebelumnya yang dilakukan oleh Risky Novendri dan rekan-rekannya dengan judul "Sentiment Analysis of YouTube Movie Trailer Comments Using Naïve Bayes" mengevaluasi kinerja model menggunakan confusion matrix. Dalam penelitian tersebut, mereka menggunakan 80% data *training* dan 20% data *testing*. Hasil evaluasi menunjukkan bahwa model mencapai akurasi sebesar 81%, dengan *precision* 74,83% dan *recall* 75,22%[11].

Penelitian lain yang dilakukan oleh Ahmad Harun dan rekan-rekannya dengan judul "Analisa Sentimen Opini Publik Tentang Vaksinasi Covid-19 di Indonesia Menggunakan Naïve Bayes dan Decision Tree" juga menunjukkan hasil yang menarik. Dalam penelitian tersebut, akurasi model yang diperoleh oleh model Naïve Bayes Classifier mencapai 100%, sedangkan akurasi model yang diperoleh oleh Decision Tree hanya mencapai 50,39%[12].

3. METODE PENELITIAN

Metode penelitian ini membahas langkah-langkah penelitian yang diambil penulis. Penulis menggunakan metode *text mining* dengan model algoritma naïve bayes classifier untuk klasifikasi sentimen twitter dalam penelitian ini. Metode ini mencakup beberapa langkah, mulai dari *crawling*, *cleaning preprocessing*, *translate*, *labeling*, dan klasifikasi bayes. Penelitian dimulai dengan

pengumpulan kumpulan data yang diperoleh dari twitter melalui teknik *crawling* data menggunakan API twitter. Selanjutnya, data akan melewati tahap *cleaning* dan *preprocessing*, kemudian data akan diubah ke dalam bahasa inggris, kemudian pelabelan akan dilakukan otomatis oleh pustaka NLP yaitu 'TextBlob' menjadi tiga label: positif, negatif, dan netral[13]. Setelah itu, data akan diklasifikasikan dengan model naïve bayes untuk memprediksi sentimen dari data teks tersebut lalu di evaluasi hasilnya.

3.1. Crawling

Data yang digunakan dalam penelitian ini berasal dari cuitan-cuitan *tweet* di aplikasi Twitter. Data yang diperoleh melalui teknik pengambilan data dari web. Jumlah data yang terkumpul sebanyak 255 *rows* data *tweet* dengan kata kunci "film agak laen" dan "agak laen" dari rentang waktu 9 maret sampai 17 maret 2024 lalu diseleksi menjadi 195 *rows* yang akan diproses.

3.2. Cleaning

Proses *cleaning* data melibatkan penghapusan data yang tidak lengkap atau kosong[14]. Proses pembersihan berguna untuk mengecek keakuratan dan kebersihan data analisis.

3.2.1. Case Folding

Case folding merupakan proses mengubah huruf kapital dalam sebuah dataset menjadi huruf kecil. Metode ini umumnya digunakan dalam pengolahan teks untuk menyamakan format huruf agar lebih mudah dianalisis.

3.3. Preprocessing

Data yang didapatkan dari hasil proses *crawling* masih dalam bentuk mentah. Melalui tahap *preprocessing* teks, data-data tersebut perlu diubah ke dalam format yang dapat dipahami. Proses ini melibatkan langkah-langkah seperti normalisasi data, *stopword*, tokenisasi, serta *stemming*[15].

3.3.1. Normalisasi

Proses normalisasi bertujuan untuk memperbaiki kata-kata yang mengandung elemen singkatan, tidak baku, atau kesalahan ketik (*typo*).

3.3.2. Stopword

Kata-kata sambung yang dianggap tidak relevan dan tidak berdampak pada sentimen dalam kalimat akan dihilangkan dari analisis.

3.3.3. Tokenisasi

Pada proses tokenisasi, kalimat akan dipecah menjadi unit kata atau token yang dipisahkan oleh spasi atau tanda baca, seperti koma.

3.3.4. Stemming

Proses tersebut berfungsi mengurangi kata yang mengandung imbuhan di awalan, sisipan, akhiran, atau kombinasi imbuhan.

3.4. Translate

Setelah melalui tahapan preprocessing data akan dilanjutkan ke tahap translate data yang berfungsi untuk menerjemahkan teks ke dalam bahasa inggris agar bisa di analisis oleh klasifikasi bayes.

3.5. Pelabelan

Setelah data dikumpulkan, data akan dilabeli berdasarkan tweet yang telah dikumpulkan. Setelah data dikumpulkan, setiap sentimen diberi label dengan tiga kategori: positif, negatif, dan netral. Tweet akan diberi labeling positif jika mengandung kalimat yang positif, sebaliknya jika mengandung kalimat negatif akan diberi label negatif, dan jika kalimat mengandung kalimat yang samar-samar, maka akan dilabeli sebagai sentimen netral[16].

3.6. Klasifikasi Bayes

Proses labeling data juga akan diklasifikasikan dengan model naive bayes classifier untuk memprediksi sentimen dari data teks tersebut yang akan dilabeli positif, netral atau negatif[17].

3.7. Evaluasi Model

Pada proses evaluasi akan menguji kinerja model naive bayes classifier dengan menghitung akurasi, precision, recall, dan f1-score. Salah satu bentuk pengujian Cross Validation adalah K-Fold Cross Validation, yang bertujuan untuk mengevaluasi kinerja suatu metode algoritma dengan membagi sampel data secara acak menjadi K kelompok, kemudian menguji setiap kelompok dengan menggunakan data dari kelompok lainnya[18].

4. HASIL DAN PEMBAHASAN

4.1. Crawling

Metode scraping digunakan pada cuitan tweet dan menggunakan bantuan bahasa pemrograman Python di Google Colaboratory untuk mengambil data dalam penelitian ini. Data yang berhasil diambil mencakup 255 baris tweet dengan kata kunci "film agak laen" dan "agak laen" dari rentang waktu 9 Maret hingga 17 Maret 2024. Data tersebut kemudian disaring menjadi 195 baris yang akan diproses dan disimpan dalam format CSV.



Gambar 1. Hasil Scraping Data

4.2. Cleaning

Cleaning data adalah proses untuk mengurangi jumlah dan kompleksitas data yang diproses. Pembersihan data dilakukan untuk menghilangkan dataset yang tidak nyambung, salah, kosong, double, dan tidak bersih sebab keberadaannya dapat mengurangi mutu atau ketepatan hasil. Pada tahap ini, ditemukan 1 data duplikat yang kemudian dihapus lalu tersisa 194 data tweet yang akan dilanjutkan ke tahap yang selanjutnya. Kemudian juga, pada tahap ini dilakukan pembersihan karakter '@' yang biasanya untuk menghilangkan nama pengguna Twitter dan non- alphanumeric serta masalah spasi berlebihan yang mungkin muncul setelah pembersihan teks sebelumnya.

4.2.1. Case Folding

Case folding merupakan langkah pembersihan data teks dengan merubah abjad kapital berubah menjadi huruf kecil. Hal ini biasanya dilakukan untuk menghindari perbedaan dalam pengejaan kata yang sama karena perbedaan kapitalisasi.

Tabel 1. Contoh Hasil Case Folding

Table with 2 columns: Sebelum and Sesudah. Row 1: Minimal sekali seumur hidup lu harus nonton film Agak Laen di Bioskop vs minimal sekali seumur hidup lu harus nonton film agak laen di bioskop.

4.3. Preprocessing

Semua komentar dalam dataset diproses selama tahap preprocessing untuk menghapus kata yang tidak berguna melibatkan beberapa tahapan seperti normalisasi, stopword, tokenisasi, dan stemming.

4.3.1. Normalisasi

Normalisasi merupakan langkah untuk memperbaiki data yang mengandung singkatan, tidak baku, atau kesalahan penulisan. Hal ini membantu memastikan konsistensi dalam teks dan mempermudah pemrosesan selanjutnya.

Tabel 2. Hasil Contoh Normalisasi

Table with 2 columns: Sebelum and Sesudah. Row 1: minimal sekali seumur hidup lu harus nonton film agak laen di bioskop vs minimal sekali seumur hidup kamu harus nonton film agak laen di bioskop.

4.3.2. Stopword

Kata-kata sambung yang tidak berdampak pada proses analisis akan dihilangkan. Hal ini bertujuan untuk menyederhanakan kalimat dan meningkatkan kejelasan pesan yang disampaikan.

Tabel 3. Hasil Contoh Stopword

Table with 2 columns: Sebelum and Sesudah. Row 1: minimal sekali seumur hidup kamu harus nonton film agak laen di bioskop vs minimal sekali seumur hidup kamu harus nonton film agak laen bioskop.

4.3.3. Tokenisasi

Pada proses tokenisasi kalimat, teks akan dipisahkan menjadi unit-unit kata, yang disebut token, yang dibatasi oleh spasi atau tanda baca seperti koma. Hal ini dilakukan untuk memudahkan analisis dan pemrosesan teks lebih lanjut.

Tabel 4. Hasil Contoh Tokenisasi

Sebelum	Sesudah
minimal sekali seumur hidup kamu harus nonton film agak laen bioskop	minimal, sekali, seumur, hidup, kamu, harus, nonton, film, agak, laen, bioskop

4.3.4. Stemming

Proses ini akan mengurangi kata-kata yang mengandung awalan, sisipan, akhiran, atau kombinasi imbuhan. Tujuannya adalah untuk menyatukan kata-kata yang memiliki akar kata yang sama, sehingga mempermudah analisis dan pemrosesan teks.

Tabel 5. Contoh Hasil Stemming

Sebelum	Sesudah
minimal sekali seumur hidup kamu harus nonton film agak laen di bioskop	minimal sekali seumur hidup kamu harus nonton film agak laen bioskop

4.4. Translate

Setelah melalui proses *preprocessing*, data akan diarahkan ke tahap *translate* untuk menerjemahkan teks ke bahasa Inggris. Langkah ini diperlukan agar teks dapat dianalisis oleh model klasifikasi Bayes.

Tabel 6. Contoh Hasil Translate

Sebelum	Sesudah
minimal sekali seumur hidup kamu harus nonton film agak laen bioskop	at least once in your life you have to watch agak laen film at the cinema

4.5. Pelabelan

Setelah melalui tahapan *cleaning*, *preprocessing* dan penerjemahan data selesai, setiap sentimen akan dilabeli menggunakan *tweet* yang diperoleh dari hasil *scraping* data berdasarkan kategori masing-masing *tweet* yang dapat dilihat pada gambar 2. hasil pelabelannya.

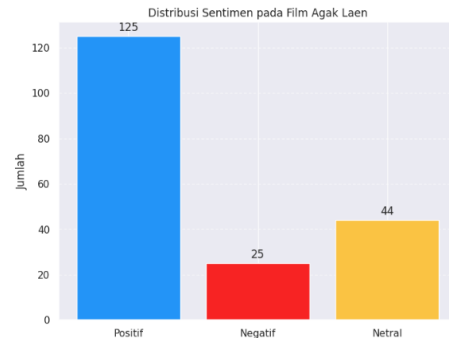
full_text	english_tweet	label
mvs dua poster kirim punya teori konspirasi ja...	mvs two posters send have conspiracy theories ...	Positif
laen seperti di hujung napas lima hari akhir h...	laen like at the end of the last five days the...	Positif
avatar agak laen idraang jegel katabene dion s...	avatar rather laen idraang jegel katabene dion...	Netral
mvs knetz notice film laen kirim apa mrka nger...	mvs knetz notice film laen send what mrka unde...	Netral
four film kuasa layar bioskop hari exhuma dua ...	four movies power screen cinema day exhuma two...	Positif
...
film agak laen empat aku nonton masih lucu ajah	ive watched agak laen film four times and its ...	Positif
biasa nonton film luar jadi ki lihat film indo...	i usually watch foreign films so when i see an...	Negatif
film agak laen sempurna pekan sahabat buat min...	a slightly agak laen film thats perfect for yo...	Positif
rekor banyak nonton film aku cuma kkn des tar...	record for watching a lot of films i just dont...	Positif
hebat lah agak laen berapa juta tonton saja br...	its great i guess how many millions just watch...	Positif

Gambar 2. Hasil Pelabelan Data

Tabel 7. Contoh Hasil Klasifikasi Bayes

full text	english tweet	label	klasifikasi bayes
minimal sekali seumur hidup kamu harus nonton film agak laen bioskop	at least once in your life you have to watch agak laen film at the cinema	negatif	positif

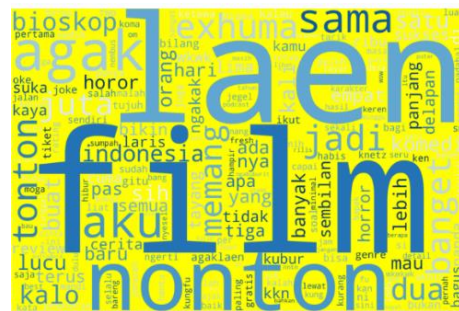
Dari hasil pelabelan di atas, ditemukan sebanyak 125 data *tweet* positif, 25 data *tweet* negatif dan 44 data *tweet* netral.



Gambar 3. Grafik Distribusi Sentimen Hasil Pelabelan

4.5.1. Visualisasi Word Cloud

Untuk menemukan tren kata paling banyak terlihat baik cuitan positif, negative, maupun netral, visualisasi *cloud* kata akan menganalisis 194 data cuitan.



Gambar 4. Visualisasi Word Cloud

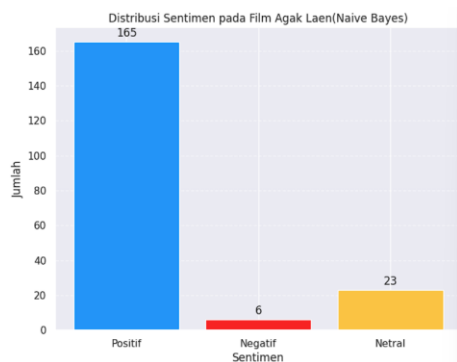
Hasil dari pemrosesan *cloud* kata, terlihat kata-kata seperti laen, film, nonton, agak, bioskop, lucu, dan sebagainya adalah yang paling sering muncul.

4.6. Klasifikasi Bayes

Selanjutnya, data yang telah dilabeli tadi diklasifikasikan dengan *naïve bayes classifier* menggunakan pustaka 'textblob.classifiers' di mana data dibagi berdasarkan label sentimen dari dataset. Setiap kategori dibagi menjadi dua set secara acak, masing-masing dengan setengah jumlah data dari kategori tersebut.

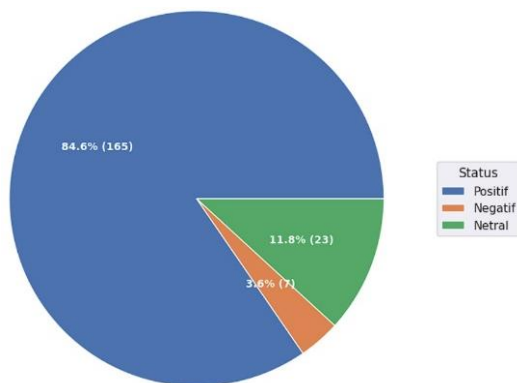
Dengan pembuatan model 'NaiveBayesClassifier' yang dilatih menggunakan data yang telah dibagi sebelumnya. Setiap data dilabeli sesuai dengan sentimen yang diberikan. Setelah model terlatih, setiap *tweet* dari *dataset* akan diklasifikasikan menggunakan model yang telah dibuat.

Dari hasil klasifikasi bayes, diperoleh ditemukan sebanyak 165 data *tweet* positif, 6 data *tweet* negatif dan 23 data *tweet* netral.



Gambar 5. Grafik Distribusi Sentimen Hasil Klasifikasi Bayes

Berdasarkan hasil klasifikasi bayes juga dapat dihasilkan presentase sentimen cuitan *tweet* para pengguna Twitter terhadap film Agak Laen yaitu dari 194 dataset, 84.6% diantaranya berstatus positif, 3.6% diantaranya berstatus negatif dan 11.8% diantaranya berstatus netral.



Gambar 6. Grafik Hasil Presentase Sentimen Naive bayes

4.7. Evaluasi Model

Diketahui hasil tes akurasi model klasifikasi bayes menunjukkan bahwa model memiliki akurasi sebesar 0,778350514639175 atau 78%. Akurasi diukur dengan nilai antara 0 hingga 1, di mana semakin tinggi nilainya, semakin baik performa model dalam mengklasifikasikan data uji.

Pada tahap ini, model klasifikasi dievaluasi menggunakan *K-Fold Cross Validator* dengan nilai $k=5$. Data dibagi menjadi 5 bagian/*fold* yang sama besar secara acak. Selanjutnya, model dilatih dan dievaluasi menggunakan lima konfigurasi yang berbeda, di mana masing-masing iterasi menggunakan satu lipatan sebagai data *test* dan lipatan lainnya sebagai data *train*. Pada setiap iterasi, model melakukan prediksi pada data uji dan membandingkan prediksi dengan label yang sebenarnya untuk menghasilkan metrik evaluasi, seperti *precision*,

recall, dan *f1-score*. Setelah semua iterasi selesai, nilai rata-rata dari metrik evaluasi dihitung untuk memberikan gambaran umum tentang performa model.

Tabel 8. Hasil Evaluasi Model

Akurasi	Precision	Recall	F1-Score
0,78 atau 78%	85,51%	85,27%	82,71%

Precision: 85.51%

Recall: 85.57%

F1-score: 82.71%

Gambar 7. Hasil Evaluasi Model

Hasil menunjukkan bahwa model klasifikasi bayes yang dievaluasi memiliki performa yang cukup baik dalam mengklasifikasi data. Diperoleh *precision* sebesar 85.51% berarti sekitar 85.51% dari *tweet* yang diprediksi positif oleh model memang benar-benar positif. Kemudian diperoleh juga *recall* sebesar 85.57% menunjukkan bahwa sekitar 85.57% dari *tweet* yang sebenarnya positif telah diidentifikasi dengan benar oleh model. Lalu, diperoleh *f1-score* sebesar 82.71% yang menunjukkan keseimbangan antara *precision* dan *recall*.

5. KESIMPULAN DAN SARAN

Dalam penelitian ini, analisis sentimen film Agak Laen dilakukan menggunakan klasifikasi Naive Bayes dengan *text mining*, karena metode ini efektif dan cepat untuk mengklasifikasikan teks. Hasil pengolahan data menunjukkan bahwa data cuitan tentang film Agak Laen dapat diklasifikasikan dengan baik menjadi label positif, negatif dan netral. Berdasarkan hasil pengujian, ditemukan bahwa penggunaan metode *text mining* dalam analisis klasifikasi sentimen *review* film menghasilkan akurasi sebesar 78%, *precision* sebesar 85.51%, *recall* sebesar 85.57%, dan *f-1 score* sebesar 82.71% yang menunjukkan bahwa model memiliki tingkat keakuratan yang memuaskan dalam mengklasifikasikan sentimen ulasan film. Menunjukkan penggunaan algoritma klasifikasi Naive Bayes Classifier mempunyai potensi yang besar untuk melakukan analisis sentimen terhadap ulasan film di media sosial Twitter. Hasil ini memberikan kontribusi yang berharga dalam pemahaman terhadap respons dan penilaian pengguna terhadap film "Agak Laen", serta dapat menjadi landasan bagi pengembangan penelitian dan aplikasi lebih lanjut dalam bidang analisis sentimen dan *text mining*.

Maka saran yang dapat diterapkan bagi pengembangan lebih lanjut seperti, memperhatikan dan memperbaiki proses pembersihan data sebagai ahap kritis dalam memastikan kualitas data yang digunakan untuk analisis. Selanjutnya, pengembang dapat memperluas analisis sentimen untuk mencakup lebih dari sekadar klasifikasi positif, negatif, dan

netral, yaitu mengembangkan analisis sentimen untuk mengidentifikasi emosi yang terkandung dalam ulasan, seperti kebahagiaan, kesedihan, kekecewaan, atau kegembiraan, dapat memberikan wawasan yang lebih dalam tentang bagaimana penonton merespons film. Serta memperluas *dataset* melalui penambahan data dapat memperluas hal dan jenis yang akan dianalisis oleh algoritma. Hal ini dapat menghasilkan model yang lebih kuat dan lebih representatif dalam mengklasifikasikan sentimen ulasan film.

DAFTAR PUSTAKA

- [1] R. Pratama, "Film Doctor Strange in The Multiverse of Madness (2022) Sebagai Budaya Populer: Sebuah Perspektif Penonton Indonesia," *Endogami J. Ilm. Kaji. Antropol.*, vol. 5, no. 2, pp. 86–87, 2022, [Online]. Available: <https://ejournal.undip.ac.id/index.php/endogami/article/view/46553>
- [2] N. S. Fathullah, Y. A. Sari, and P. P. Adikara, "Analisis Sentimen Terhadap Rating dan Ulasan Film dengan menggunakan Metode Klasifikasi Naïve Bayes dengan Fitur Lexicon-Based," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 4, no. 2, pp. 590–593, 2020.
- [3] M. T. Razaq, D. Nurjanah, and H. Nurrahmi, "Analisis Sentimen Review Film Menggunakan Naive Bayes Classifier dengan Fitur TF-IDF," *e-Proceeding Eng.*, vol. 10, no. 2, pp. 1698–1712, 2023.
- [4] D. Alita and A. R. Isnain, "Pendeteksian Sarkasme pada Proses Analisis Sentimen Menggunakan Random Forest Classifier," *J. Komputasi*, vol. 8, no. 2, pp. 50–58, 2020, doi: 10.23960/komputasi.v8i2.2615.
- [5] N. Fibriyanti Arminda, N. Sulistiyowati, and T. Nur Padilah, "Implementasi Algoritma Multinomial Naive Bayes Pada Analisis Sentimen Terhadap Ulasan Pengguna Aplikasi Brimo," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 7, no. 3, pp. 1817–1822, 2023, doi: 10.36040/jati.v7i3.7012.
- [6] H. Rahab, A. Zitouni, and M. Djoudi, "SANA: Sentiment analysis on newspapers comments in Algeria," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 33, no. 7, pp. 899–907, 2021, doi: 10.1016/j.jksuci.2019.04.012.
- [7] R. Nanda, E. Haerani, S. K. Gusti, and S. Ramadhani, "Klasifikasi Berita Menggunakan Metode Support Vector Machine," *J. Nas. Komputasi dan Teknol. Inf.*, vol. 5, no. 2, pp. 269–278, 2022, doi: 10.32672/jnkti.v5i2.4193.
- [8] K. Anwar, "Analisa sentimen Pengguna Instagram Di Indonesia Pada Review Smartphone Menggunakan Naive Bayes," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 2, no. 4, pp. 148–155, 2022, doi: 10.30865/klik.v2i4.315.
- [9] A. R. Rinaldi, T. Informatika, and K. D. Timur, "PENERAPAN METODE NAÏVE BAYES CLASSIFIER PADA ANALISIS SENTIMEN APLIKASI GOPAY," vol. 8, no. 1, pp. 346–353, 2024.
- [10] S. A. R. Rizaldi, S. Alam, and I. Kurniawan, "Analisis Sentimen Pengguna Aplikasi JMO (Jamsostek Mobile) Pada Google Play Store Menggunakan Metode Naive Bayes," *STORAGE J. Ilm. Tek. dan Ilmu Komput.*, vol. 2, no. 3, pp. 109–117, 2023, doi: 10.55123/storage.v2i3.2334.
- [11] R. Novendri, A. S. Callista, D. N. Pratama, and C. E. Puspita, "Sentiment Analysis of YouTube Movie Trailer Comments Using Naïve Bayes," *Bull. Comput. Sci. Electr. Eng.*, vol. 1, no. 1, pp. 26–32, 2020, doi: 10.25008/bcsee.v1i1.5.
- [12] A. Harun and D. Putri Ananda, "Analisa Sentimen Opini Publik Tentang Vaksinasi Covid-19 di Indonesia Menggunakan Naive bayes dan Decision Tree," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 1, no. 1, pp. 58–64, 2021, doi: 10.57152/malcom.v1i1.63.
- [13] R. Azhar, A. Surahman, and C. Juliane, "Analisis Sentimen Terhadap Cryptocurrency Berbasis Python TextBlob Menggunakan Algoritma Naïve Bayes," *J. Sains Komput. Inform. (J-SAKTI)*, vol. 6, no. 1, pp. 267–281, 2022.
- [14] Y. Astuti, I. R. Wulandari, A. R. Putra, and N. Kharomadhona, "Naïve Bayes untuk Prediksi Tingkat Pemahaman Kuliah Online Terhadap Mata Kuliah Algoritma Struktur Data," *J. Edukasi dan Penelit. Inform.*, vol. 8, no. 1, p. 28, 2022, doi: 10.26418/jp.v8i1.48848.
- [15] Yuyun, Nurul Hidayah, and Supriadi Sahibu, "Algoritma Multinomial Naïve Bayes Untuk Klasifikasi Sentimen Pemerintah Terhadap Penanganan Covid-19 Menggunakan Data Twitter," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 4, pp. 820–826, 2021, doi: 10.29207/resti.v5i4.3146.
- [16] Irvandi, B. Irawan, and O. Nurdiawan, "Naive Bayes Dan Wordcloud Untuk Analisis Sentimen Wisata Halal Pulau Lombok," *INFOTECH J.*, vol. 9, no. 1, pp. 236–242, 2023, doi: 10.31949/infotech.v9i1.5322.
- [17] U. Kusnia and F. Kurniawan, "Analisis Sentimen Review Aplikasi Media Berita Online Pada Google Play menggunakan Metode Algoritma Support Vector Machines (SVM) Dan Naive Bayes," *J. Keilmuan dan Apl. Tek. Inform.*, vol. 5, no. 36, pp. 22–28, 2022.
- [18] R. R. R. Arisandi, B. Warsito, and A. R. Hakim, "Aplikasi Naïve Bayes Classifier (Nbc) Pada Klasifikasi Status Gizi Balita Stunting Dengan Pengujian K-Fold Cross Validation," *J. Gaussian*, vol. 11, no. 1, pp. 130–139, 2022, doi: 10.14710/j.gauss.v11i1.33991.