

Klasifikasi Sentimen Opini terhadap Film Kartini Menggunakan Naive Bayes pada Platform X

Muhammad Al-Fajr Ramadhani*, Joko Aryanto, Iwan Hartadi Tri Untoro, Agus Sujarwadi

Fakultas Sains & Teknologi, Program Studi Informatika, Universitas Teknologi Yogyakarta, Yogyakarta, Indonesia

Email: ^{1,*}itsmerandom688@gmail.com, ²joko.aryanto@uty.ac.id, ³iwan.hartadi@uty.ac.id, ⁴agus.sujarwadi@uty.ac.id

Email Penulis Korespondensi: itsmerandom688@gmail.com

Submitted: 06/11/2024; Accepted: 28/11/2024; Published: 29/11/2024

Abstrak—Film merupakan media audiovisual yang tidak hanya berfungsi sebagai hiburan, tetapi juga dapat memberikan informasi dan menyentuh emosi penontonnya. Dalam perkembangan era digital, masyarakat cenderung menggunakan media sosial untuk mencari film yang sedang populer atau menjadi tren. Salah satu platform yang sering digunakan untuk tujuan tersebut adalah Twitter, yang pada Juli 2023 berganti nama menjadi 'X'. Penelitian ini membahas tentang analisis sentimen terhadap film Kartini di platform media sosial X, yang bercerita tentang perjuangan R.A. Kartini dalam memperjuangkan hak-hak perempuan Indonesia. Film ini menjadi topik pembicaraan dengan berbagai opini, baik positif maupun negatif yang beredar di platform tersebut. Banyaknya sentimen masyarakat seringkali menimbulkan kebingungan bagi calon penonton dalam menentukan apakah film tersebut layak untuk ditonton. Variasi pandangan terhadap film ini bisa memengaruhi persepsi penonton, yang pada akhirnya berisiko menciptakan pengalaman menonton yang mengecewakan dan membuat penonton merasa bahwa waktu yang dihabiskan tidak berharga. Oleh karena itu, peneliti bertujuan untuk melakukan klasifikasi sentimen pengguna platform X terhadap film Kartini dengan menggunakan metode Naive Bayes Classifier, yang bekerja berdasarkan Teorema Bayes untuk memprediksi probabilitas suatu kejadian berdasarkan data-data sebelumnya. Metode ini dipilih karena kemampuannya dalam mengklasifikasikan sentimen ke dalam kategori positif atau negatif secara efisien. Hasil sementara pada data testing dengan menggunakan algoritma naive bayes dan bahasa pemrograman python didapatkan hasil yang sangat baik. Hasil akurasi terbaik dari dataset emosi yang didapatkan adalah 98% dan akurasi terbaik dari dataset sentimen yang didapatkan adalah 86%.

Kata Kunci: Analisis Sentimen; Naive Bayes; Film Kartini; Platform X; Python

Abstract—Film is an audiovisual media that not only serves as entertainment, but can also provide information and touch the emotions of its audience. In the development of the digital era, people tend to use social media to search for films that are popular or trending. One platform that is often used for this purpose is Twitter, which in July 2023 changed its name to 'X'. This research discusses the sentiment analysis of the film Kartini on the X social media platform, which tells the story of R.A. Kartini's struggle for the rights of Indonesian women. The film became a topic of conversation with various opinions, both positive and negative, circulating on the platform. The amount of public sentiment often creates confusion for potential viewers in determining whether the film is worth watching. These varying views on the film can affect viewers' perceptions, which in turn risks creating a disappointing viewing experience and making viewers feel that the time spent was not worthwhile. Therefore, the researcher aims to classify the sentiment of platform X users towards the film Kartini using the Naive Bayes Classifier method, which works based on Bayes' Theorem to predict the probability of an event based on previous data. This method was chosen due to its ability to efficiently classify sentiment into positive or negative categories. Temporary results on testing data using the naive bayes algorithm and python programming language obtained very good results. The best accuracy result of the emotion dataset obtained is 98% and the best accuracy of the sentiment dataset obtained is 86%.

Keywords: Sentiment Analysis; Naive Bayes; Kartini Film; Platform X; Python

1. PENDAHULUAN

Film merupakan media audiovisual yang tidak hanya berfungsi sebagai hiburan, tetapi juga dapat memberikan informasi serta menyentuh emosi penonton. Menurut Hiawan Pratista (2008), film menggabungkan dua unsur penting, yaitu naratif dan sinematik. Unsur naratif berkaitan dengan tema cerita, sementara unsur sinematik mencakup alur yang runtun dari awal hingga akhir. Aspek naratif membentuk struktur kronologis cerita, dengan fokus pada elemen-elemen seperti karakter, dialog, plot, latar, dan tema, sedangkan elemen sinematik meliputi aspek teknis seperti *mise-en-scène*, sinematografi, penyuntingan, dan audio [1]. Dalam perkembangan era digital, masyarakat cenderung menggunakan media sosial untuk mencari film yang sedang populer atau trending. Saat ini, media sosial banyak dimanfaatkan sebagai sarana komunikasi, penyebaran informasi, edukasi, hiburan, serta platform untuk kegiatan pemasaran [2]. Media sosial juga dapat memberikan manfaat dalam konteks perfilman, salah satunya adalah sebagai media untuk memberikan penilaian dan penjelasan atau ulasan tentang film yang disebut dengan *movie review* [3]. Salah satu platform yang sering digunakan untuk keperluan ini adalah Twitter, yang pada bulan Juli 2023 berganti nama menjadi "X".

Media sosial berfungsi sebagai ruang interaksi virtual yang memungkinkan pengguna saling terhubung dan membangun relasi sosial, didukung oleh beragam fitur yang memenuhi kebutuhan mereka [4]. Platform X memungkinkan pengguna untuk berbagi teks, gambar, video, serta berinteraksi melalui like, komentar, repost, dan pesan langsung. Platform ini telah menjadi tempat utama bagi masyarakat untuk berbagi pengalaman, termasuk rekomendasi film. Dengan film, masyarakat bisa memahami berbagai fenomena yang pernah terjadi atau sedang berlangsung, baik berdasarkan kisah nyata maupun fiksi [5]. Salah satu film yang banyak

dibicarakan adalah Kartini, yang mengisahkan perjuangan tokoh emansipasi perempuan Indonesia, R.A. Kartini, dalam memperjuangkan hak-hak perempuan, khususnya dalam bidang pendidikan. Perjuangan Kartini dalam film ini menyoroti kesetaraan gender dan emansipasi, sesuatu yang jarang dibahas dalam buku-buku sejarah [6]. Keberanian Kartini tidak lepas dari peran sang kakak, Sostrokartono (Reza Rahadian), yang memberinya berbagai buku untuk dibaca agar mampu mewujudkan impiannya untuk menyetarakan hak perempuan.

Film Kartini mendapat banyak perhatian publik, baik dalam bentuk ulasan positif maupun negatif. Meskipun demikian, banyaknya sentimen yang beragam terhadap film ini di platform X seringkali membingungkan calon penonton. Mereka mungkin merasa ragu apakah film tersebut layak ditonton atau tidak, mengingat variasi opini yang ada. Kondisi ini dapat memengaruhi pengalaman menonton, bahkan berpotensi menimbulkan kekecewaan atau kerugian jika ekspektasi penonton tidak terpenuhi. Untuk mengatasi permasalahan ini, diperlukan analisis yang sistematis terhadap sentimen-sentimen tersebut. Salah satu metode yang efektif dan sering digunakan dalam analisis sentimen adalah Naive Bayes Classifier. Metode ini bekerja berdasarkan Teorema Bayes yang memprediksi peluang suatu peristiwa berdasarkan data sebelumnya. Dalam konteks analisis sentimen, metode ini mampu mengklasifikasikan opini publik menjadi sentimen positif atau negatif secara efisien.

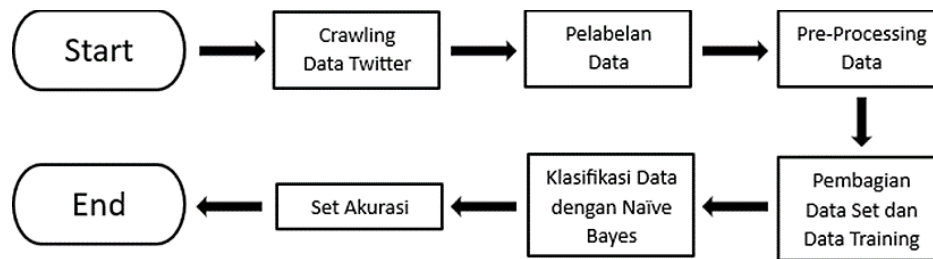
Penelitian terdahulu oleh Amrullah dkk. (2020) membahas tentang suatu ulasan yang berisi opini atau pandangan penonton terhadap suatu karya film. Penelitian ini menggunakan 1.400 data opini film berbahasa Inggris, terdiri dari 700 opini positif dan 700 opini negatif. Sebanyak 1.000 data (500 opini positif dan 500 opini negatif) digunakan sebagai data latih, sementara 400 data sisanya digunakan sebagai data uji. Dua pengujian dilakukan dengan membagi data uji dan data latih dalam dua skenario berbeda [7]. Penelitian yang dilakukan oleh Putri dkk. (2022) membahas tentang analisis sentimen menggunakan metode klasifikasi Naive Bayes terhadap 550 data latih dari postingan Instagram dengan kategori komentar positif (105), negatif (43), netral (288), sangat positif (37), dan sangat negatif (36). Analisis sentimen ini menghasilkan akurasi sebesar 60,7073% [8]. Kemudian penelitian dari Nurtikasari dkk. (2022) meneliti tentang sentimen masyarakat pada platform twitter terhadap kasus film Ngeri-nger Sedap. Pelabelan data Twitter dilakukan secara manual dengan pemberian sentimen positif, negatif, dan netral pada dataset mentah di Microsoft Excel, sebelum masuk tahap preprocessing, meliputi transformasi, tokenisasi, dan penyaringan. Setelah proses transformasi selesai, pembobotan tf-idf digunakan untuk mengukur frekuensi kemunculan kata. Data kemudian diklasifikasikan menggunakan Algoritma Naive Bayes. Pengujian menggunakan confusion matrix dengan tools Orange menghasilkan rata-rata akurasi 65%, presisi 67%, recall 65%, dan klasifikasi netral sebesar 83% [9]. Selanjutnya penelitian oleh Kurniawan dan Akbar (2024) membahas tentang identifikasi gangguan tidur pada anak, sistem ini menggunakan algoritma Naive Bayes sebagai metode klasifikasi probabilistik sederhana yang berbasis pembelajaran mesin dan menghitung probabilitas. Sistem ini menganalisis 31 gejala gangguan tidur pada anak, mencakup jenis-jenis gangguan seperti Sleep Apnea, Insomnia, Narkolepsi, Enuresis, Night Terror, Nightmare, dan Sleepwalking. Berdasarkan pengujian yang dilakukan menggunakan 20 data kasus dari para ahli, sistem ini mencapai akurasi sebesar 95% [10]. Dan yang terakhir adalah penelitian oleh Effendi dkk (2024) ini membahas tentang sentimen masyarakat twitter terhadap presiden 2024 yaitu Anis, Prabowo dan Ganjar dengan metode klasifikasi algoritma Naive Bayes. Dari perbandingan hasil analisis ketiga kandidat calon presiden, nilai akurasi Anies Baswedan lebih rendah dibandingkan dua kandidat lainnya, karena Anies Baswedan menerima lebih banyak komentar negatif. Anies Baswedan memperoleh nilai akurasi sebesar 67,23%, Prabowo Subianto sebesar 83,42%, dan Ganjar Pranowo sebesar 88,15%. Jumlah data yang tersedia turut mempengaruhi hasil analisis sentimen; semakin banyak data yang digunakan, semakin baik pula akurasi yang dihasilkan [11].

Penelitian ini bertujuan untuk mengklasifikasikan sentimen opini pengguna di platform X terhadap film Kartini dengan menggunakan metode Naive Bayes. Hasil penelitian yang didapatkan adalah data testing dengan menggunakan algoritma Naive Bayes dan bahasa pemrograman python mencapai akurasi terbaik dari dataset emosi yang didapatkan adalah 98% dan akurasi terbaik dari dataset sentimen yang didapatkan adalah 86%. Diharapkan hasil tersebut dapat membantu masyarakat dalam memahami kualitas film berdasarkan sentimen yang tersebar di platform tersebut, serta memudahkan dalam pengambilan keputusan sebelum menonton.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Pada Gambar 1. merupakan tahapan penelitian untuk membangun sebuah sistem analisis atau klasifikasi sentimen opini terhadap film menggunakan algoritma Naive Bayes. Tahapan penelitian memiliki peran krusial dalam menjamin kelancaran proses secara terstruktur, sehingga mampu menghasilkan data yang tepat dan terpercaya demi tercapainya tujuan penelitian secara optimal. Pengguna bertanggung jawab dalam mengkonfigurasi dan mengelola seluruh infrastruktur sistem perangkat lunak untuk proyek. Tugas utama pengguna meliputi pengaturan sistem dan input data.



Gambar 1. Tahapan Penelitian

Proses ini dilakukan menggunakan Google Colab sebagai perangkat lunak. Data akan dipisahkan menjadi dua bagian, yaitu data uji (*testing*) dan data latih (*training*), untuk menerapkan model yang dipilih. Algoritma yang digunakan dalam analisis ini adalah Naive Bayes. Setelah pemrosesan selesai, laporan hasil akan dikirimkan kembali kepada pengguna.

2.1.1 Crawling Data Twitter

Pengumpulan data otomatis, yang dikenal sebagai *crawling data*, adalah proses yang dimulai dengan mengakses URL target atau situs web yang telah ditentukan, dalam hal ini adalah Platform X yang akan dikunjungi untuk melakukan proses *scraping* atau pengambilan data dari halaman yang ditentukan [12].

2.1.2 Pelabelan Data

Pelabelan data adalah proses menambahkan label, kategori, atau kode tertentu pada setiap entitas atau elemen data untuk mengidentifikasi atau mengklasifikasikan data berdasarkan karakteristik atau atribut yang spesifik [13]. Dengan menambahkan label pada setiap elemen data, model dapat belajar untuk mengenali pola tertentu antara input dan output, sehingga meningkatkan kemampuan untuk melakukan prediksi atau klasifikasi.

2.1.3 Pre-Processing Data

Preprocessing data merupakan langkah awal sebelum proses klasifikasi, yang berfungsi untuk membersihkan data dengan menghapus atau mengubah karakter non-alfabet serta kata-kata yang tidak relevan, sehingga data lebih siap dan optimal untuk digunakan dalam proses klasifikasi [14]. Proses transformasi data mentah ke dalam format yang lebih terstruktur dan siap dianalisis sangatlah penting, mengingat data mentah sering kali tidak teratur serta mengandung kesalahan, nilai kosong, error, dan informasi yang kurang relevan.

2.1.4 Pembagian Data

Pembagian dataset pada machine learning sangat penting karena memungkinkan model belajar lebih efektif, mencegah overfitting, menyempurnakan hyperparameter, serta mengukur performa dan generalisasi model secara objektif di data yang belum pernah dilihat sebelumnya.

2.1.5 Klasifikasi Data

Naive Bayes adalah algoritma klasifikasi yang menggunakan Teorema Bayes untuk memprediksi probabilitas kejadian di masa depan berdasarkan data sebelumnya. Metode ini memandang probabilitas sebagai tingkat keyakinan terhadap kebenaran suatu pernyataan dan menggunakan probabilitas bersyarat untuk membuat keputusan klasifikasi, dengan formulasi dasar:

$$P(C | X) = \frac{P(X|C) \cdot P(C)}{P(X)} \tag{1}$$

Di sini, $P(C|X)$ adalah probabilitas kelas C mengingat fitur X , $P(X|C)$ adalah probabilitas data fitur mengingat kelas C , $P(C)$ adalah probabilitas awal dari kelas, dan $P(X)$ adalah probabilitas data fitur. Untuk atribut numerik, probabilitas dalam Naive Bayes dinyatakan sebagai berikut:

$$P(Y | X) = \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{(x_i - \mu_j)^2}{2\sigma_j^2}} \tag{2}$$

Di mana x_i adalah nilai atribut, σ_j adalah standar deviasi, dan μ_j adalah rata-rata dari atribut tersebut dalam kelas tertentu.

2.1.6 Set Akurasi

Akurasi dalam Naive Bayes adalah ukuran seberapa efektif model dalam memprediksi kelas pada data uji, dihitung sebagai persentase dari prediksi yang benar dibandingkan dengan total data. Metrik ini krusial untuk menilai kinerja model, menyempurnakan parameter, dan membandingkan berbagai model. Namun, penting

untuk mempertimbangkan distribusi kelas dalam data, karena ketidakseimbangan dapat memengaruhi hasil akurasi.

2.2 Pengumpulan Data

Sumber data yang digunakan adalah data sekunder, yaitu informasi yang sudah ada dan dikumpulkan secara sengaja oleh peneliti untuk melengkapi kebutuhan penelitian. Data berasal dari platform X (2023–2024) terkait komentar positif dan negatif tentang film, yang diperoleh melalui *crawling* data menggunakan kode program *tweet harvest* dengan kata kunci “Kartini”.

2.3 Data Penelitian

Dalam penelitian ini, data yang digunakan mencakup seluruh sentimen masyarakat Indonesia yang memberikan komentar tentang film Kartini dalam periode waktu dari tahun 2017 hingga 2024. Data sentimen tersebut dikumpulkan melalui metode *crawling*, yang merupakan teknik untuk mengambil data secara otomatis dari platform X. Dengan menggunakan data yang luas dan beragam ini maka akan mendapatkan wawasan yang lebih mendalam mengenai persepsi publik terhadap film Kartini.

conversat_created_at	favorite_count	full_text	id_str	image_url_in_reply_to_lang	location	quote_count	reply_count	retweet_count	tweet_url	user
1.8E+18 Wed Jun 1	0	Aku sejarah selain gurunya enak dan menyenangkan. Lewat pelajaran ini aku lebih	1.8E+18		Indonesia	0	0	0	https://x.c.1.	
1.8E+18 Mon Jun 1	0	film biografi tokoh Indonesia selain soekarno kartini sama bj habibie apalagi ya	1.8E+18			0	0	0	https://x.c.1.	
1.8E+18 Sun Jun 16	14	Digitalkin bu KampanyalarA+ sadece bizde. 12 ay Sabit Sporun ya+ldA+zA+ + 12	1.8E+18	https://pbs.twimg.cctr		0	3	4	https://x.c.8.	
1.8E+18 Sat Jun 15	3	MILBBANC yang dia jadi main/supporter roles: srumulat ada di prime video one ni	1.8E+18	MILBBANC	abin har	0	1	0	https://x.c.1.	
1.8E+18 Sat Jun 15	11	Digitalkin bu KampanyalarA+ sadece bizde. 12 ay Sabit Sporun ya+ldA+zA+ + 12	1.8E+18	https://pbs.twimg.cctr		0	3	1	https://x.c.8.	
1.8E+18 Fri Jun 14	1	Digitalkin bu KampanyalarA+ sadece bizde. 12 ay Sabit Sporun ya+ldA+zA+ + 12	1.8E+18	https://pbs.twimg.cctr		0	2	0	https://x.c.8.	
1.8E+18 Wed Jun 1	0	Buat mengembangkan au Himawari [mingmin x wonmina] aku menonton ulang film	1.8E+18		saturn.	0	1	0	https://x.c.1.	
1.8E+18 Tue Jun 11	0	abis nonton film kartini tapi ke distract mulu ama kebaya nya cakep bgt pleasee f	1.8E+18		hell	0	0	0	https://x.c.1.	
1.8E+18 Tue Jun 11	0	nonton bareng film Kartini yang akan dilaksanakan pada: Sabtu 25 Mei 2024 9:00:00	1.8E+18		bemkfmkr in	0	1	0	https://x.c.6.	
1.8E+18 Fri Jun 07	0	@moviementes nder coba liat kartini 2017 pas awal film aja aku udah nangis keingi	1.8E+18		moviemer in	0	0	0	https://x.c.8.	
1.8E+18 Fri Jun 07	0	Special Ulang Tahun @keirashabira nobar Film Dua Surga Dalam Cintaku 8 Juni 202	1.8E+18	https://pbs.twimg.ccin	Indonesia	0	0	0	https://x.c.1.	
1.8E+18 Fri Jun 07	1	@robertrony @therealDISastr @officialpilarez @shelladaisha @AdiniaWirasti @ac	1.8E+18		robertroni in	0	0	0	https://x.c.1.	
1.8E+18 Fri Jun 07	12	Sebelum pengumuman pemenang bagi-bagi T-shirt merchandise film yg saya perne	1.8E+18	https://pbs.twimg.ccin	New York	3	20	4	https://x.c.33	
1.8E+18 Thu Jun 06	0	Jatuh cinta berkali-kali sama film dan buku KARTINI j	1.8E+18		Jawa Timu	0	0	0	https://x.c.1.	
1.8E+18 Wed Jun 0	0	@drercardio 6-7 ya+ldA+r TV izlemiyordum. Ortanca ASocuAYum futbol kursuna gidi	1.8E+18		leynnnn_gtr	0	1	0	https://x.c.9.	
1.8E+18 Wed Jun 0	0	@oceanizblue Ya film kartini yg main dian sastro kalo gak salah	1.8E+18		oceanizbl in	0	1	0	https://x.c.1.	
1.8E+18 Wed Jun 0	0	@jipaveda woww film kartini itu yang mnaa	1.8E+18		jipaveda in	0	1	0	https://x.c.1.	
1.8E+18 Wed Jun 0	0	1.5cm 2. Magic hour (film ya bukan jeketi) 3. Cheer up (kdrama) 4. Kartini 5. Merah j	1.8E+18		Bogor,Jaw	0	1	0	https://x.c.1.	
1.8E+18 Wed Jun 0	0	@Eastcrookd Semangat <3 cobain nonton film Sang Penari juga deh. Ini bukan fi	1.8E+18		Eastcrook in	0	0	0	https://x.c.8.	
1.8E+18 Thu May 3	0	Tidak ada yang tau siapa Kartini sebelum film ini dibuat: https://t.co/WIDd@wiPdY	1.8E+18	https://pbs.twimg.ccin	Indonesia	0	0	0	https://x.c.1.	

Gambar 2. Data Mentah

Pada Gambar 2 merupakan contoh dari dataset yang akan digunakan dalam penelitian ini, di mana nantinya hanya akan menggunakan satu atribut, yaitu *full_text*. Jumlah data yang diperoleh sekitar 500 data.

full_text
Pernah nonton film kartini nya dian sastro. Mama ku malah tidur di tengah-tengah film
Makanya pak @ganjarpranowo stop nonton film porno yah bersama istri karena orang yang blunder itu orang yang doyan film porno.
@Andi00583950 @OposisiCerdas @HusinShihab Bang Husin GK ada niat barter film bokep dgn @ganjarpranowo Bokepers kagak pantas bicara politik karena SDH tidak punya harga dirinya lagi
@Lisaamartatara6 @AfifFuadS Penyuka film porno harus didukung sama penyuka film porno dok fitrahnya begitu makanya kagak heran kyai NU byk memperkosa santriatnya.
@IndraJa13989548 @puspita3_eka @rasulkita1 Rumah produksi kok tidak baik FTV sinetron dan filn film lainnya dibuat rumah produksi. Salahnya dimana kak?? Heloo
Dinbudpar Rembang Sosialisasikan Film Animasi RA Kartini #beritarembang #nurfmrembang https://t.co/3MuJ6n7z6z
kardeşim okulda bi soruyu çözdüğü için öğretmenim ödül olarak çizgi film karakterlerinden bir kart seçtirip hediye etmiş. kardeşim bana hediye etmek için Tinkerbelle kartını seçmiş. tatlılığını anlatmaya kelime bulamıyorum
1. Nonton Film Thailand Baifern & Mario Maurer sambil nunggu jam Pramuka 2. Ikut Lomba Baca Puisi Hari Kartini 3. Ikut Karnaval & Nari Topeng Kaliwungu 4. Ikut Jalan Candil 5. Ikut Pramuka 6. Lebih fokus ke ujian praktek bikin mini film dari pada belajar buat UN/Try Out
buat kakak yang ngejoki review film kartini sama aku please do reach out me through DM yaa!! WA aku lagi bermasalah filenya aku kirim besok pagi karena tugasnya on process! pls semoga kakaknya ngecek akun twit aku ya

Gambar 3. Data Dengan Atribut Full_Text

Pada Gambar 3 merupakan contoh dari dataset yang telah menggunakan atribut *full_text*. Atribut ini memungkinkan pencarian efisien dalam dokumen panjang, analisis mendalam terhadap konten teks, dan penyimpanan data besar, serta meningkatkan performa melalui indeksasi, fleksibilitas penggunaan, dan integrasi dengan pemrosesan bahasa alami (NLP).

3. HASIL DAN PEMBAHASAN

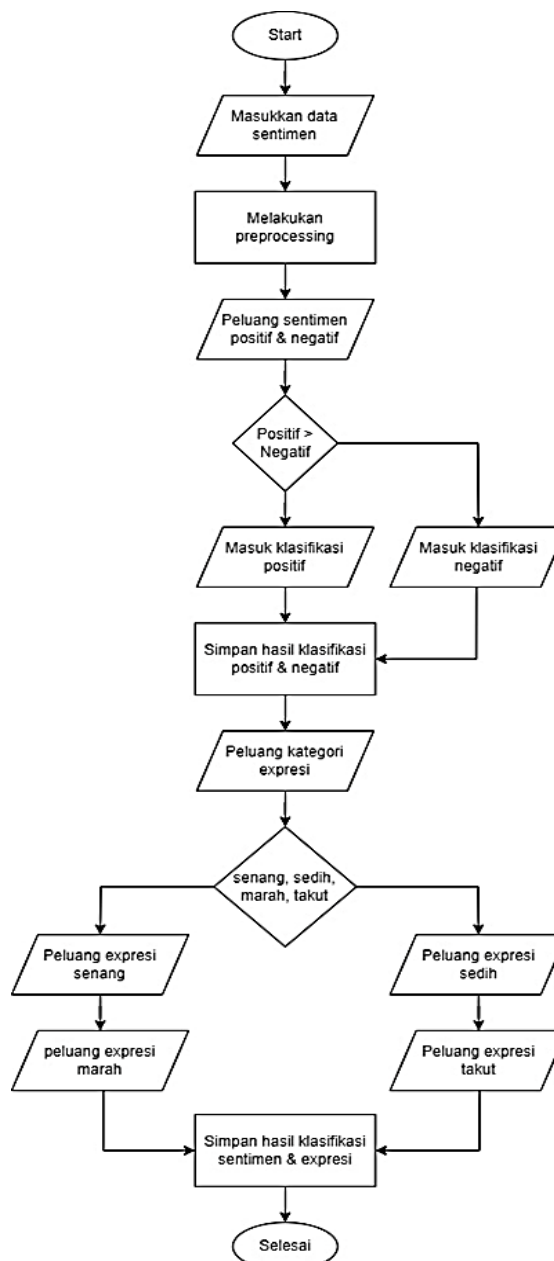
Pada bagian hasil dan pembahasan aplikasi mengevaluasi performa, fitur, serta efektivitas aplikasi dalam memenuhi kebutuhan yang dirancang, meliputi antarmuka, kemudahan penggunaan, dan ketepatan output. Selain itu tahapan ini juga berfungsi untuk mengidentifikasi kelebihan dan kekurangan sistem seperti kecepatan, keakuratan, bug atau area yang memerlukan peningkatan untuk menilai sejauh mana aplikasi memenuhi kebutuhan pengguna dan potensi pengembangan lebih lanjut.

3.1 Perancangan Sistem

Perancangan logika merupakan proses penyusunan struktur sistem atau program untuk memperlihatkan bagaimana data diproses dan bagaimana keputusan diambil sesuai kondisi yang ada. Dalam perancangan ini, penyajian dilakukan melalui representasi *Flowchart*, *Activity Diagram* dan Diagram Alir Data (DAD) digunakan untuk memvisualisasikan alur proses, interaksi antara pengguna dan sistem secara jelas dan terstruktur [15].

3.1.1 Flowchart

Flowchart adalah representasi visual yang menunjukkan langkah-langkah dan urutan prosedur dalam sebuah program [16]. Flowchart memudahkan analisis proses dengan menggunakan notasi standar untuk menggambarkan aktivitas, kondisi, dan alur logika, serta dapat diterapkan di berbagai bidang seperti pengembangan perangkat lunak, perencanaan bisnis, dan manajemen proyek [17].

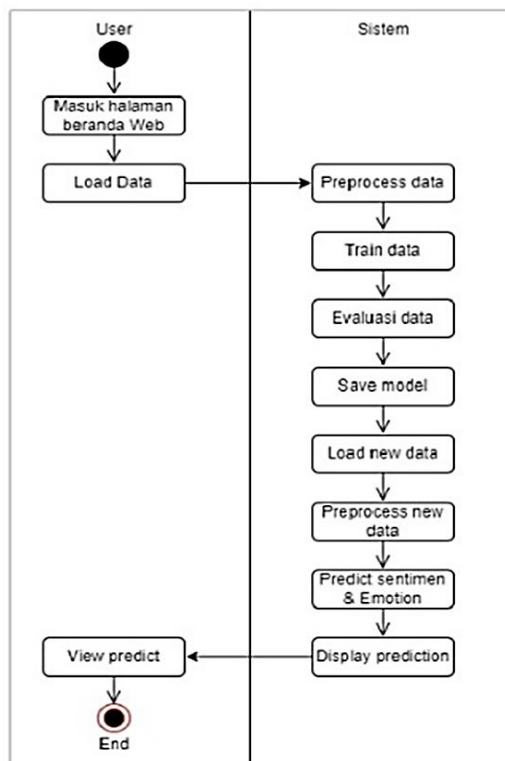


Gambar 4. Flowchart

Pada Gambar 4 menunjukkan alur sistem dalam diagram flowchart. Pada penelitian ini, proses dimulai dengan pengumpulan data sentimen, yang kemudian diolah sesuai prosedur penelitian. Selanjutnya, data tersebut diklasifikasikan menggunakan algoritma Naïve Bayes. Hasil akhir berupa klasifikasi sentimen dan ekspresi kemudian dapat disimpan.

3.1.2 Activity Diagram

Activity diagram menggambarkan alur kerja sistem dalam perangkat lunak berdasarkan rancangan Use Case, dengan fokus pada aktivitas yang dijalankan sistem, bukan tindakan dari aktor [18]. Penjelasan urutan proses pada diagram aktivitas membantu mempermudah pemahaman tentang alur kerja dalam sistem [19].

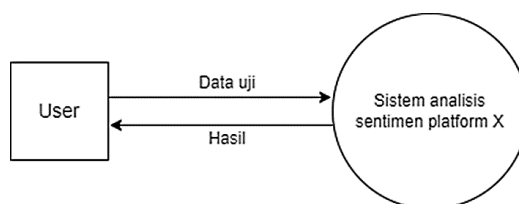


Gambar 5. Activity Diagram

Pada Gambar 5 merupakan gambaran suatu proses atau aktivitas pengguna pada saat menggunakan sistem. Pengguna masuk halaman web sistem dan memasukkan data, lalu sistem akan melakukan proses klasifikasi data berdasarkan prosedur yang telah ditentukan. Selanjutnya hasil akhir berupa prediksi yang diinginkan berdasarkan data yang telah dimasukkan akan ditampilkan kembali kepada pengguna melalui halaman web.

3.1.3 Diagram Alir Data (DAD)

Diagram Alir Data (DAD) memanfaatkan notasi-notasi untuk menunjukkan alur data dalam sistem, sehingga memudahkan pemahaman sistem secara logis, terstruktur, dan jelas [20]. Diagram alir data memiliki manfaat seperti memvisualisasikan proses, mengidentifikasi masalah, meningkatkan komunikasi, mendokumentasikan sistem, menganalisis kebutuhan, serta mendukung pengembangan dan desain yang lebih efisien, sekaligus mendeteksi redundansi dan meningkatkan kualitas sistem secara keseluruhan.



Gambar 6. Diagram Alir Data (DAD)

Diagram ini memperlihatkan alur data dari satu proses ke proses lainnya, bagaimana data disimpan, serta bagaimana data diterima atau dihasilkan oleh sistem. Pada Gambar 6 dijelaskan bahwa pengguna menginput data uji yang diterima oleh sistem untuk dianalisis. Setelah analisis selesai, hasil data langsung diterima oleh pengguna.

3.2 Implementasi

Hasil implementasi dapat dijalankan menggunakan browser Google Chrome dengan mengunjungi sistem yang akan dibuat pada Google Colab menggunakan bahasa pemrograman python. Dibawah ini merupakan langkah dalam pembuatan sistem analisis sentimen opini film kartini pada platform x menggunakan metode naive bayes.

```
# Import library yang diperlukan
import pandas as pd
import numpy as np
import re
import nltk
from nltk.tokenize import WhitespaceTokenizer
from nltk.stem import WordNetLemmatizer
from sklearn.model_selection import train_test_split
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.pipeline import make_pipeline
from sklearn.naive_bayes import MultinomialNB
from sklearn.metrics import accuracy_score
```

Gambar 7. Import Library

Pada Gambar 7 menampilkan proses impor library yang dibutuhkan dalam penelitian ini. Library adalah kumpulan modul yang sudah ditulis sebelumnya dan dapat digunakan oleh program lain untuk mempercepat pengembangan aplikasi. Library menyediakan berbagai fungsi yang bisa diakses dan dimanfaatkan oleh pengguna untuk menangani tugas tertentu tanpa harus membuat kode dari awal.

```
Data Train_Kartini_PP(550).csv
Twitter_Emotion.csv

# Membaca data dari Google Drive untuk Dataset Emotion dan Sentiment
data_emotion = pd.read_csv('/content/drive/My Drive/Data Baru/Twitter_Emotion.csv')
data_sentiment = pd.read_csv('/content/drive/My Drive/Data Baru/Data Train_Kartini_PP(550).csv')
```

Gambar 8. Input Dataset Yang Digunakan

Bagian ini berfungsi untuk membaca dua dataset dari Google Drive penulis, yang akan digunakan untuk menganalisis sentimen dan emosi pada data tweet. Data tersebut ditampilkan pada Gambar 8.

```
# Membersihkan teks pada kolom 'full_text' dari dataset baru
uji['clean_text'] = uji['full_text'].apply(clean_text_new)

# Menghapus kata berhenti dari setiap kalimat dalam kolom 'clean_text' dataset baru
uji['clean_text'] = uji['clean_text'].apply(remove_stopwords)

# Melakukan lemmatisasi pada kolom 'clean_text' dataset baru
uji['clean_text'] = uji['clean_text'].apply(lemmatize_text)
```

Gambar 9. Preprocessing

Pada Gambar 9 menunjukkan tahap preprocessing yang dilakukan menggunakan fitur *clean_text*, *remove_stopword*, dan *lemmatize_text*. Fungsi *clean_text* berperan dalam membersihkan kata atau kalimat yang memiliki anomali, seperti menghapus tag HTML dan elemen lainnya. Kemudian *remove_stopword* berfungsi untuk menghapus kata-kata umum yang tidak memiliki makna signifikan. Sementara itu, *lemmatize_text* berfungsi untuk mengurangi variasi kata yang berbeda tetapi memiliki arti dasar yang sama, sehingga kompleksitas analisis teks dapat diminimalkan.

```
# Mengambil nilai teks yang telah dibersihkan sebagai fitur untuk dataset baru
X_uji = uji['clean_text'].values

# Melakukan prediksi pada dataset baru menggunakan model Naive Bayes yang telah dilatih
y_pred_uji_emotion = model_emotion.predict(X_uji)
y_pred_uji_sentiment = model_sentiment.predict(X_uji)

# Menambahkan hasil prediksi ke dalam DataFrame dataset baru
uji['predicted_emotion'] = y_pred_uji_emotion
uji['predicted_sentiment'] = y_pred_uji_sentiment

# Menampilkan hasil prediksi
uji
```

Gambar 10. Tahap Prediksi

Setelah melakukan preprocessing data pada tahap sebelumnya, langkah selanjutnya adalah memulai tahap prediksi Gambar 10 pada dataset baru menggunakan Naive Bayes yang telah dilatih, serta menambahkan hasil prediksi ke dalam DataFrame dataset baru. Hasil prediksi tersebut dapat dilihat pada Gambar 11. berikut.

	full_text	clean_text	predicted_emotion	predicted_sentiment
0	Pernah nonton film kartini nya dian sastro. Ma...	pernah nonton film kartini nya dian sastro mam...	sedih	positive
1	Makanya pak @ganjarpranowo stop nonton film po...	makanya pak ganjarpranowo stop nonton film por...	senang	negative
2	@Andi00583950 @OposisiCerdas @HusinShihab Bang...	andi00583950 oposisicerdas husinshihab bang hu...	marah	positive
3	@Lisaamartatara6 @AfifFuadS Penyuka film porno...	lisaamartatara6 afiffuads penyuka film porno h...	marah	negative
4	@IndraJa13989548 @puspita3_eka @rasulkita1 Rum...	indraja13989548 puspita3eka rasulkita1 rumah p...	marah	positive
...
160	paling gampang contohnya ya film Kartini aja. ...	paling gampang contohnya ya film kartini aja t...	marah	positive
161	@moviemenfes Percuma panjang lebar sok analisa...	moviemenfes percuma panjang lebar sok analisa ...	marah	positive
162	Dalam film Kartini Jenny sangat luwes menjadi ...	dalam film kartini jenny sangat luwes menjadi ...	sedih	positive
163	Film Kartini hampir membawa Jenny memenangkan ...	film kartini hampir membawa jenny memenangkan ...	sedih	positive

Gambar 11. Hasil Prediksi

```
import joblib

# Memuat model emosi dari file
model_emotion = joblib.load('/content/drive/My Drive/Data Baru/model_emotion.pkl')

# Memuat model sentimen dari file
model_sentiment = joblib.load('/content/drive/My Drive/Data Baru/model_sentiment.pkl')
```

Gambar 12. Library Joblib

Joblib digunakan untuk menyimpan dan memuat model dalam sistem, sehingga model dapat dimuat kembali tanpa perlu dilatih ulang, memungkinkan pengguna untuk langsung menggunakan model tersebut untuk prediksi atau analisis lebih lanjut.

```
import re
import pandas as pd
from sklearn.naive_bayes import MultinomialNB
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.pipeline import Pipeline
```

Gambar 13. Library Tambahan

Pada Gambar 13 *Regular expressions* (re) digunakan untuk membersihkan teks dari karakter khusus, sementara Pandas memfasilitasi pembacaan dan manipulasi data dalam berbagai format, seperti CSV dan Excel. MultinomialNB adalah algoritma Naive Bayes yang cocok untuk klasifikasi teks dengan data diskrit. CountVectorizer mengubah teks menjadi representasi vektor berdasarkan frekuensi kata. Penggunaan pipeline memungkinkan preprocessing teks dan pelatihan model dilakukan secara efisien dalam satu langkah.

```
# Teks yang ingin diuji dalam bahasa Indonesia
text_to_test = "film yang sangat bagus namun masih ada yang kurang"

# Melakukan prediksi emosi dan sentimen dari teks yang diberikan
predicted_emotion, predicted_sentiment = predict_emotion_sentiment(text_to_test)

# Menampilkan hasil prediksi
print("Text:", text_to_test)
print("Predicted Emotion:", predicted_emotion)
print("Predicted Sentiment:", predicted_sentiment)
```

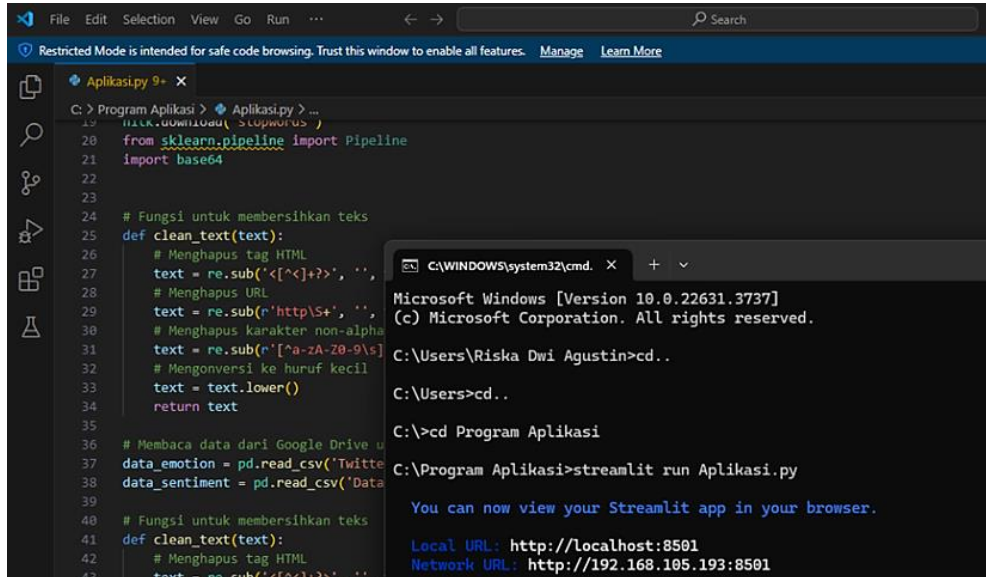
Text: film yang sangat bagus namun masih ada yang kurang
Predicted Emotion: sedih
Predicted Sentiment: positive

Gambar 14. Pengujian

Pada Gambar 14 menunjukkan pengujian teks menggunakan bahasa Indonesia dengan input "film yang sangat bagus namun masih ada yang kurang." Sistem kemudian menganalisis emosi dan sentimen secara akurat, dan hasilnya menampilkan teks yang telah diketik, dengan *predicted emotion* menunjukkan emosi "sedih" dan *predicted sentiment* menunjukkan sentimen positif.

3.3 Hasil

Sistem analisis pada platform X yang telah dirancang sebelumnya dapat diimplementasikan dan dijalankan melalui command prompt. Sistem ini dibangun menggunakan framework Streamlit, yang menawarkan tampilan sederhana dan dirancang untuk memudahkan pengembangan aplikasi berbasis machine learning.



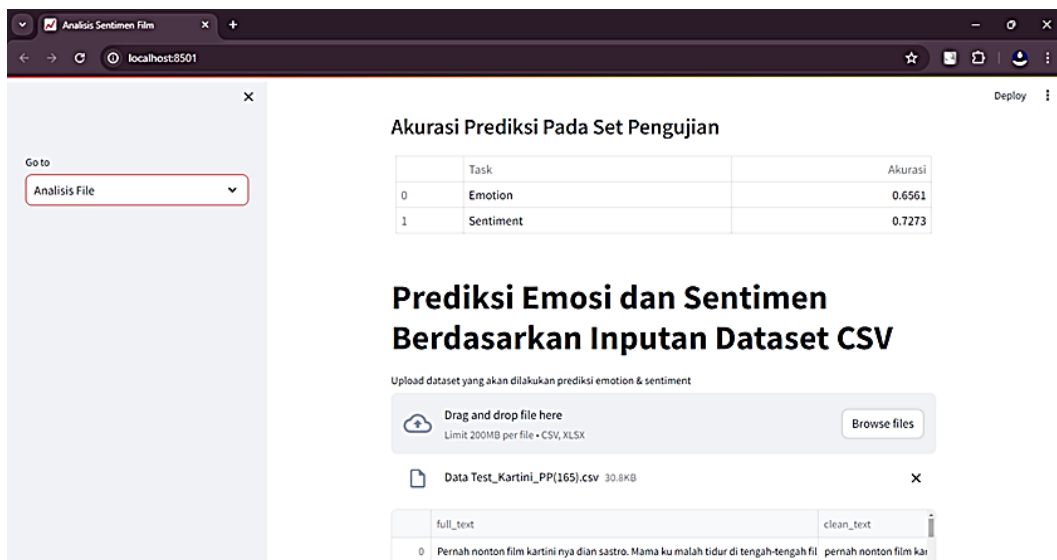
Gambar 15. Perintah Menjalankan Sistem

Penjelasan dalam Gambar 15 dimulai dengan membuka file Python yang telah dibuat, kemudian menjalankan command prompt dan memilih folder yang berisi file sistem. Setelah itu, ketikkan “streamlit run Aplikasi.py” untuk memulai.



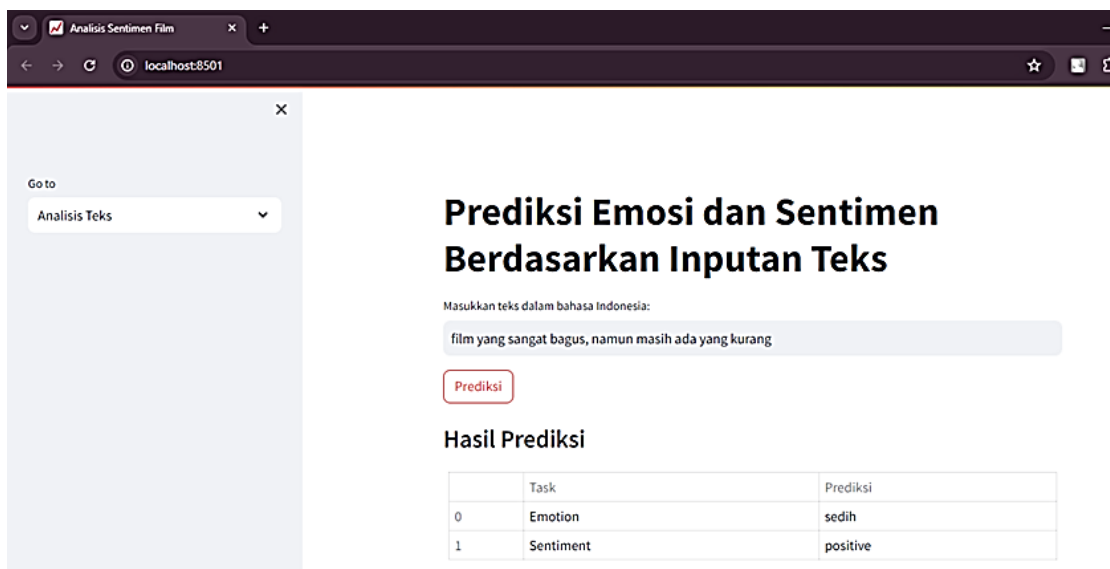
Gambar 16. Dashboard

Setelah mengikuti langkah untuk menjalankan sistem, pada Gambar 16. pengguna secara otomatis akan diarahkan ke Google Chrome, di mana beranda atau menu utama dari aplikasi analisis sentimen film pada platform X akan ditampilkan.



Gambar 17. Menu Analisis Teks

Pada menu Gambar 17 ini, ditampilkan akurasi prediksi pada set pengujian, serta fitur untuk mengunggah dataset CSV, yang memungkinkan pengguna melihat data yang telah dibersihkan. Selain itu, menu ini juga menampilkan prediksi emosi dan prediksi sentimen.



Gambar 18. Menu Analisis File

Kemudian pada menu Gambar 18 tersebut, pengguna dapat memasukkan teks pada kolom yang tersedia, misalnya dengan mengetik "film yang sangat bagus, namun masih ada yang kurang," dan menekan tombol prediksi. Hasil prediksi akan ditampilkan berdasarkan teks yang telah diinputkan.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa algoritma Naive Bayes dapat diterapkan secara efektif untuk analisis sentimen pada komentar di platform X. Data yang digunakan terdiri dari dua jenis, yaitu dataset emosi dan dataset sentimen. Melalui pengujian yang dilakukan, algoritma Naive Bayes menghasilkan akurasi yang memuaskan, yaitu 98% untuk dataset emosi dan 86% untuk dataset sentimen. Penerapan analisis ini dilakukan menggunakan bahasa pemrograman Python, yang memungkinkan proses analisis menjadi lebih efisien. Dengan hasil yang diperoleh, penelitian ini diharapkan dapat menciptakan lingkungan informasi yang lebih positif, memberikan panduan berharga bagi pengguna, serta meningkatkan kepuasan masyarakat terhadap pilihan film yang mereka buat. Selain itu, berdasarkan hasil analisis, terdapat saran untuk penelitian selanjutnya agar dapat meningkatkan jumlah data yang digunakan, yang diharapkan dapat memberikan akurasi yang lebih baik pada tahap preprocessing. Penelitian berikutnya juga disarankan untuk menambahkan kategori netral, sehingga analisis dapat lebih jelas dalam menggambarkan opini yang tidak bersifat positif maupun negatif. Dengan langkah-langkah ini, diharapkan analisis sentimen dapat terus berkembang dan memberikan manfaat yang lebih besar di masa mendatang.

REFERENCES

- [1] N. D. Thaheer and V. D. Adiprabowo, "Analisis Naratif Dalam Film Singot," *Tonil: Jurnal Kajian Sastra, Teater dan Sinema*, vol. 21, no. 1, pp. 15–27, 2024.
- [2] S. Saepudin, S. Widiastuti, and C. Irawan, "Sentiment Analysis Of Social Media Platform Reviews Using The Naïve Bayes Classifier Algorithm," *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, vol. 12, no. 2, 2023.
- [3] A. L. Afriani, D. Suprayitno, and N. A. Misbah, "Pengaruh Media Online Reviews terhadap Keputusan Menonton Film," *Jurnal Penelitian Sosial Ilmu Komunikasi*, vol. 7, no. 1, 2023.
- [4] A. Khairani, "Konstruksi Hoax Melalui Fitur Media Sosial Dalam Perspektif Cyberculture," *DIALEKTIKA KOMUNIKA: Jurnal Kajian Komunikasi dan Pembangunan Daerah*, vol. 7, no. 2, 2020.
- [5] A. Putri and L. Nurhajati, "Representasi Perempuan Dalam Kungkungan Tradisi Jawa Pada Film Kartini Karya Hanung Bramantyo," *ProTVF*, vol. 4, no. 1, 2020.
- [6] M. Fahmi, "Hegemoni Kesetaraan Gender Pada Film Kartini Karya Hanung Bramantyo," *Musawa: Jurnal Studi Gender Dan Islam*, vol. 19, no. 1, 2020.
- [7] A. Z. Amrullah, A. S. Anas, and Muh. A. J. Hidayat, "Analisis Sentimen Movie Review Menggunakan Naive Bayes Classifier Dengan Seleksi Fitur Chi Square," *Jurnal Bumigora Information Technology (BITE)*, vol. 2, no. 1, 2020.

- [8] S. B. Putri, Y. N. Anisa, and N. Saputra, “Analisis Sentimen Film Kuliah Kerja Nyata (KKN) Di Desa Penari Menggunakan Metode Naive Bayes,” *Journal of Information Systems and Communication Technology (JuSiTik)*, vol. 5, no. 2, 2022.
- [9] Y. Nurtikasari, S. Alam, and T. I. Hermanto, “Analisis Sentimen Opini Masyarakat Terhadap Film Pada Platform Twitter Menggunakan Algoritma Naive Bayes,” *INSOLOGI: Jurnal Sains dan Teknologi*, vol. 1, no. 4, 2022.
- [10] B. D. Kurniawan and M. Akbar, “Sistem Pakar Diagnosa Gangguan Tidur Pada Anak Menggunakan Naive Bayes,” *Journal of Computer System and Informatics (JoSYC)*, vol. 5, no. 4, pp. 747–757, 2024.
- [11] M. M. Effendi, A. T. Zy, and A. Arwan, “Analisis Sentimen Media Sosial Twitter Terhadap Calon Presiden RI Tahun 2024 Menggunakan Klasifikasi Algoritma Naive Bayes,” *Journal of Computer System and Informatics (JoSYC)*, vol. 5, no. 3, pp. 739–746, 2024.
- [12] L. Lathifah, E. Handoyo, and Y. A. A. Soetrisno, “Sistem Crawling Data Instrumen Akreditasi Berbasis Selenium Dan Pandas,” *Transient: Jurnal Ilmiah Teknik Elektro*, vol. 10, no. 1, pp. 84–91, 2021.
- [13] T. Gori, A. Sunyoto, and H. Al Fatta, “Preprocessing Data dan Klasifikasi untuk Prediksi Kinerja Akademik Siswa,” *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 11, no. 1, 2024.
- [14] I. Romli and R. F. P. Dewi, “Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Klasifikasi Penyakit ISPA,” *Indonesian Journal of Business Intelligence (IJUBI)*, vol. 4, no. 1, 2021.
- [15] B. O. Lubis, B. Santoso, R. T. Yunandar, and B. A. Wahid, “Desain Sistem Informasi Pemesanan Paket Wisata dengan Metode Framework Application of System Thinking (FAST),” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 3, 2024.
- [16] A. Sutanti, M. K. MZ, Mustika, and P. Damayanti, “Rancang Bangun Aplikasi Perpustakaan Keliling Menggunakan Pendekatan Terstruktur,” *Jurnal Ilmiah Dan Informatika (KOMPUTA)*, vol. 9, no. 1, 2020.
- [17] T. E. Ardiansyah, E. H. Priyatna, and T. Megaster, *Management Information System Dengan Pendekatan Literasi Perangkat Dan Etika Teknologi Informasi*. Purbalingga: Eureka Media Aksara, 2023.
- [18] M. N. Umam, D. Solahudin, and A. R. Rinaldi, “Pengembangan Aplikasi Perencanaan Pernikahan Berbasis Mobile,” *EKUITAS: Jurnal Teknologi Ilmu Komputer*, vol. 2, no. 1, pp. 89–94, 2023.
- [19] A. R. Putri, A. Hafizhah, F. H. Rahmah, R. Muslihhah, and Nabila Suci, “Pemodelan Diagram UML Pada Perancangan Sistem Aplikasi Konsultasi Hewan Peliharaan Berbasis Android (Studi Kasus: Alopel),” *Jurnal Ilmu Komputer Dan Bisnis (JIKB)*, vol. 12, no. 2, pp. 130–139, 2021.
- [20] Y. M. Putra and F. H. Silalahi, “Diagram Aliran Data untuk Perusahaan Kecil (Studi Kasus Pada PT. APK),” *Jurnal AKSI (Akuntansi Dan Sistem Informasi)*, 2021.