



1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.
2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

BAB II

TINJAUAN LITERATUR

Bab ini mengulas ringkasan penelitian terdahulu yang berkaitan dengan Analisis *Sentimen* terhadap ulasan pengguna aplikasi dengan memanfaatkan metode klasifikasi dalam bidang text mining. Tinjauan literatur dilakukan untuk mengetahui hasil penelitian sebelumnya serta melihat efektivitas metode yang digunakan dalam mengolah data ulasan pengguna. Berdasarkan hasil kajian dari beberapa penelitian sebelumnya, algoritma Naive Bayes sering dipilih untuk analisis sentimen karena prosesnya sederhana, memiliki waktu komputasi yang cepat, serta mampu memberikan hasil klasifikasi dengan tingkat akurasi yang baik dalam membedakan sentimen positif, negatif, dan netral. Metode ini bekerja berdasarkan konsep probabilitas dari Teorema Bayes, yaitu menghitung kemungkinan suatu dokumen atau teks termasuk ke dalam kelas tertentu berdasarkan kata-kata yang muncul di dalamnya.

2.1 Analisis *Sentimen*

Analisis *Sentimen* adalah salah satu cabang dari text mining yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini, perasaan, atau pendapat yang terdapat dalam suatu teks [6]. Analisis *Sentimen* dimanfaatkan untuk memahami bagaimana pendapat pengguna terhadap suatu produk atau layanan berdasarkan ulasan yang mereka berikan. Melalui proses ini, opini pengguna dapat dikelompokkan ke dalam sentimen positif, negatif, atau netral sehingga memudahkan proses evaluasi terhadap produk atau layanan tersebut.

Dalam praktiknya, Analisis *Sentimen* sering digunakan untuk menganalisis ulasan pengguna pada media sosial, forum diskusi, maupun platform aplikasi



digital. Hasil Analisis *Sentimen* biasanya diklasifikasikan ke dalam beberapa kategori seperti *sentimen* positif, negatif, dan netral. Penerapan analisis sentimen memungkinkan data ulasan pengguna diproses secara otomatis meskipun jumlahnya cukup banyak. Hasil analisis tersebut memberikan gambaran mengenai kecenderungan pendapat pengguna sehingga dapat dimanfaatkan sebagai bahan evaluasi untuk meningkatkan kualitas produk atau layanan.

2.2 Algoritma *Naive Bayes*

Naive Bayes adalah salah satu algoritma klasifikasi berbasis probabilitas yang menggunakan konsep Bayes' Theorem [7]. Algoritma tersebut bekerja dengan menghitung probabilitas suatu data termasuk ke dalam kelas tertentu berdasarkan fitur atau kata-kata yang terdapat dalam dokumen.

Disebut *naive* karena algoritma ini mengasumsikan bahwa setiap fitur atau kata dalam dokumen bersifat independen satu sama lain, meskipun dalam kenyataannya hubungan antar kata bisa saja saling berkaitan. Meskipun memiliki asumsi yang sederhana, algoritma *Naive Bayes* terbukti efektif dalam berbagai penelitian, terutama dalam klasifikasi data teks seperti spam detection, analisis *sentimen*, dan kategorisasi dokumen.

- a) Keunggulan algoritma *Naive Bayes* antara lain:
- b) Proses perhitungan yang sederhana dan cepat
- c) Mampu menangani data dalam jumlah besar
- d) Memberikan hasil klasifikasi teks yang cukup akurat
- e) Mudah diimplementasikan

Dalam penelitian ini, algoritma *Naive Bayes* digunakan untuk mengklasifikasikan ulasan pengguna aplikasi Dola ke dalam kategori *sentimen*



positif, negatif, dan netral berdasarkan kata-kata yang muncul dalam ulasan tersebut. Hasil klasifikasi ini diharapkan dapat memberikan gambaran mengenai persepsi pengguna terhadap aplikasi Dola.

2.3 Aplikasi Dola AI

Dola AI adalah aplikasi asisten berbasis kecerdasan buatan yang dimanfaatkan untuk memudahkan pengguna dalam mengelola jadwal, membuat pengingat kegiatan, serta mengatur aktivitas sehari-hari secara lebih terorganisir dengan memanfaatkan teknologi *Artificial Intelligence* dan pemrosesan bahasa alami. Melalui berbagai fitur yang tersedia, aplikasi ini memungkinkan pengguna untuk membuat agenda, mengatur jadwal pertemuan, serta memperoleh pengingat otomatis terhadap kegiatan yang telah dijadwalkan. Sebagai aplikasi yang tersedia pada platform *Google Play Store*, Dola AI juga menerima berbagai ulasan dan penilaian dari pengguna sesuai dengan pengalaman yang dirasakan pengguna saat menggunakan aplikasi tersebut. Ulasan ini dapat dimanfaatkan sebagai sumber data dalam penelitian *Analisis Sentimen* untuk mengetahui persepsi pengguna terhadap kualitas dan kinerja aplikasi Dola AI.

Aplikasi DOLA AI pertama kali diperkenalkan pada tahun 2023 oleh SPRING (SG) PTE, LTD, dengan nama Cici sebagai aplikasi asisten berbasis kecerdasan buatan yang memanfaatkan teknologi *Artificial Intelligence* (AI) dan *Natural Language Processing* (NLP). Aplikasi ini dikembangkan untuk membantu pengguna dalam menjawab pertanyaan, menyusun tulisan, mengelola pekerjaan, serta mendukung aktivitas belajar dan produktivitas sehari-hari [8].

Seiring dengan perkembangan fitur dan peningkatan kemampuan layanan, pada akhir tahun 2025 pengembang melakukan rebranding dengan mengubah

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

peraturan yang berlaku di Indonesia.



nama aplikasi dari Cici menjadi DOLA. Berdasarkan informasi resmi pada riwayat pembaruan aplikasi, perubahan tersebut hanya merupakan pergantian identitas aplikasi, sedangkan berbagai fitur utama tetap dipertahankan dan akan terus dikembangkan melalui pembaruan berikutnya [8].

Setelah berganti nama menjadi DOLA, aplikasi terus dikembangkan dengan penambahan berbagai fitur berbasis AI, seperti pembuatan teks, bantuan belajar, analisis dokumen, pembuatan gambar, serta asisten produktivitas. Perkembangan tersebut menjadikan DOLA sebagai salah satu aplikasi AI yang memperoleh banyak pengguna dan ulasan di Google Play Store, sehingga layak dijadikan objek dalam penelitian analisis sentiment [8].

2.4 Google Play Store sebagai Sumber Data

Google Play Store adalah layanan distribusi aplikasi berbasis Android yang dikelola oleh Google. Platform ini memudahkan pengguna dalam menemukan, mengunduh, memasang, dan memperbarui berbagai aplikasi sesuai kebutuhan. Selain itu, pengguna juga dapat memberikan rating dan ulasan sebagai bentuk umpan balik terhadap aplikasi yang digunakan., memberikan penilaian berupa rating, serta menuliskan ulasan berdasarkan pengalaman mereka saat menggunakan aplikasi tersebut. Fitur ulasan pada Google Play Store memungkinkan pengguna menyampaikan pengalaman, kritik, dan saran setelah menggunakan suatu aplikasi. Informasi tersebut dapat dimanfaatkan sebagai bahan evaluasi untuk mengetahui kualitas aplikasi dari sudut pandang pengguna. [9], sehingga dapat dimanfaatkan dalam penelitian Analisis *Sentimen* untuk mengetahui tingkat kepuasan pengguna serta mengidentifikasi berbagai permasalahan yang terdapat pada aplikasi yang digunakan.

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

peraturan yang berlaku di Indonesia.

2.5 Penelitian Terdahulu

Berikut disajikan beberapa penelitian terdahulu yang relevan dengan topik penelitian ini:

Tabel 2. 1 Penelitian Terdahulu

NO (1)	Nama Peneliti (2)	Judul Penelitian (3)	Pembahasan (4)
1.	Riki Roska Rismansyah, Aso Sudiarjo, Teuku Mufizar (2025)	Analisis <i>Sentimen</i> Ulasan Shopee pada Google <i>Play Store</i> Menggunakan Algoritma <i>Naive Bayes</i>	Penelitian ini menganalisis sentimen pengguna aplikasi Shopee berdasarkan ulasan di Google Play Store menggunakan algoritma Naive Bayes. Sebanyak 2.000 data ulasan dikumpulkan melalui teknik web scraping dan diproses menggunakan tahapan preprocessing. Hasil penelitian memperoleh akurasi sebesar 83%, dengan sentimen negatif lebih dominan dibandingkan sentimen positif. Temuan ini dapat dimanfaatkan sebagai bahan evaluasi untuk meningkatkan kualitas layanan Shopee.

Hak Cipta Dilindungi Undang-Undang

Universitas Islam Indragiri

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.

Plagiarisme juga dilarang dan dapat dikenakan sanksi. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



Tabel 2.1 Penelitian Terdahulu (Lanjutan)

NO	Nama Peneliti	Judul Penelitian	Pembahasan
(1)	(2)	(3)	(4)
2.	Divia Azarine, Nining Rahaningsih, Raditya Danar Dana (2025)	Analisis <i>Sentimen</i> Terhadap Pengguna Aplikasi <i>Jamsostek</i> <i>Mobile</i> pada Google <i>Play Store</i> Menggunakan Metode <i>Naive Bayes</i>	Penelitian tersebut mengkaji <i>sentimen</i> pengguna aplikasi JMO berdasarkan ulasan yang tersedia di Google Play Store dengan menerapkan algoritma <i>Naive Bayes</i> . Tahapan penelitian diawali dengan pengumpulan data ulasan, kemudian dilakukan <i>preprocessing</i> yang meliputi pembersihan teks, penghapusan <i>stopword</i> , <i>stemming</i> , serta pembobotan kata menggunakan metode <i>TF-IDF</i> sebelum proses klasifikasi dilakukan. Berdasarkan hasil pengujian, model memperoleh akurasi sebesar 86,06%. Distribusi <i>sentimen</i> menunjukkan bahwa ulasan positif lebih mendominasi (74,1%) dibandingkan ulasan negatif (25,9%).
3.	Tia Septiani Gumilar, Rini Astuti, Yudhistira Arie Wijaya (2024)	Analisis <i>Sentimen</i> Ulasan Aplikasi Lita di <i>Play Store</i> Menggunakan Algoritma <i>Naive</i> <i>Bayes</i>	Penelitian ini menganalisis <i>sentimen</i> ulasan pengguna aplikasi Lita di Google <i>Play Store</i> menggunakan algoritma <i>Naive Bayes</i> . Data yang digunakan sebanyak 1000 ulasan yang diproses melalui tahapan <i>preprocessing</i> seperti <i>cleaning</i> , <i>Tokenizing</i> , dan <i>stemming</i> , serta pembobotan kata menggunakan metode <i>TF-IDF</i> . Hasil penelitian menunjukkan bahwa sebagian besar ulasan pengguna <i>bersentimen</i> positif dengan tingkat akurasi model mencapai 95%.

Hak Cipta Dilindungi Undang-Undang

Universitas Islam Indragiri

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.

Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

Tabel 2.1 Penelitian Terdahulu (Lanjutan)

NO	Nama Peneliti	Judul Penelitian	Pembahasan
(1)	(2)	(3)	(4)
4.	Wahyudi, Rudi Kurniawan, Arie Yudhistira, Wijaya (2024)	Analisis <i>Sentimen</i> Pengguna Terhadap Aplikasi Blu BCA di Playstore Menggunakan Algoritma <i>Naïve Bayes</i>	Penelitian ini menganalisis sentimen pengguna aplikasi Blu BCA berdasarkan ulasan yang tersedia di Google Play Store dengan menerapkan algoritma Naive Bayes. Data ulasan diproses melalui beberapa tahapan, yaitu web scraping, preprocessing, pembobotan kata menggunakan metode TF-IDF, hingga proses klasifikasi sentimen. Pengujian model menghasilkan akurasi sebesar 85,31%, dengan nilai precision 84,55%, recall 93,69%, dan F1-score 88,88%. Hasil tersebut menunjukkan bahwa algoritma Naive Bayes mampu memberikan performa yang baik dalam mengklasifikasikan sentimen pada ulasan pengguna aplikasi Blu BCA.
5.	Gery Riansyah Ramadhan, Agus Castaka, Sugianto (2024)	Ulasan Aplikasi DANA di Google Play Store Menggunakan Algoritma <i>Naïve Bayes</i>	Penelitian ini menganalisis <i>sentimen</i> ulasan pengguna aplikasi DANA di Google Play Store menggunakan algoritma <i>Naïve Bayes</i> . Data yang digunakan sebanyak 2000 ulasan berbahasa Indonesia yang diperoleh melalui <i>web scraping</i> . Proses penelitian meliputi tahapan <i>preprocessing</i> , pembobotan kata menggunakan <i>TF-IDF</i> , klasifikasi, dan evaluasi menggunakan <i>Confusion Matrix</i> . Hasil penelitian menunjukkan model memiliki performa yang baik dengan akurasi 86%, <i>precision</i> 86%, <i>recall</i> 98%, dan <i>F1-score</i> 92%.

Hak Cipta Dilindungi Undang-Undang

Universitas Islam Indragiri

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.

4. Plagiarisme juga dilarang dan dapat dikenakan sanksi.

5. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



Tabel 2.1 Penelitian Terdahulu (Lanjutan)

NO	Nama Peneliti	Judul Penelitian	Pembahasan
(1)	(2)	(3)	(4)
6.	Rema Sekar Nuraini, Dwi Krisbiantoro, Mohammad Imron, Pungkas Subarkah (2025)	Analisis <i>Sentimen</i> Pengguna Aplikasi Agoda di Google <i>Play Store</i> Menggunakan Algoritma <i>Naive Bayes</i>	Penelitian ini menganalisis ulasan pengguna aplikasi Agoda di Google <i>Play Store</i> menggunakan algoritma <i>Naive Bayes</i> dengan 7.885 data hasil scraping. Proses analisis meliputi tahapan <i>preprocessing</i> , pelabelan menerapkan InSet Lexicon, serta pembobotan <i>TF-IDF</i> sebelum dilakukan klasifikasi. Hasil penelitian menunjukkan akurasi terbaik sebesar 91,9%, sehingga <i>Naive Bayes</i> dinilai efektif dalam mengklasifikasikan <i>sentimen</i> pengguna secara akurat.
7.	Suci Nurhikmah, Romi Ramadani, Gandung Triyono (2025)	Analisis <i>Sentimen</i> pada Ulasan Aplikasi Wondr di <i>Play Store</i> dengan Metode <i>Naive Bayes</i>	Penelitian ini mengkaji <i>sentimen</i> pengguna aplikasi Wondr melalui ulasan yang tersedia di Google <i>Play Store</i> . Proses klasifikasi dilakukan dengan menerapkan algoritma <i>Naive Bayes</i> yang diawali dengan tahapan <i>preprocessing</i> dan pembobotan kata sebelum data dianalisis. Hasil pengujian memperoleh akurasi sebesar 82%, meskipun kemampuan model dalam mengenali <i>sentimen</i> netral masih perlu ditingkatkan. Temuan tersebut menunjukkan bahwa analisis <i>sentimen</i> dapat dimanfaatkan untuk memberikan gambaran mengenai tanggapan pengguna terhadap layanan mobile banking Wondr.

Hak Cipta Dilindungi Undang-Undang

Universitas Islam Indragiri

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia. Plagiarisme juga dilarang dan dapat dikenakan sanksi.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



Tabel 2.1 Penelitian Terdahulu (Lanjutan)

NO	Nama Peneliti	Judul Penelitian	Pembahasan
(1)	(2)	(3)	(4)
8.	Rachmawati Oktaria Mardiyanto, Kusrini, dan Ferry Wahyu Wibowo (2023)	Analisis <i>Sentimen</i> Pengguna Aplikasi Bank Syariah Indonesia Menggunakan Algoritma <i>Support Vector Machine</i> (SVM)	Penelitian ini menganalisis ulasan pengguna aplikasi BSI Mobile yang diambil dari <i>Google Play Store</i> . Data yang digunakan berjumlah 13.568 ulasan yang telah melalui proses <i>preprocessing</i> seperti <i>case folding</i> , <i>filtering</i> , <i>stopword removal</i> , <i>tokenization</i> , dan <i>stemming</i> . Metode yang digunakan adalah algoritma <i>Support Vector Machine</i> (SVM) dengan pembagian data 70% data training dan 30% data testing. Hasil pengujian menunjukkan bahwa metode SVM mampu mengklasifikasikan <i>sentimen</i> pengguna dengan nilai akurasi sebesar 85,87%.
9.	Hamzah Naufal Zuhdi, Budi Prasetyo (2025)	Analisis <i>Sentimen</i> pada Ulasan Aplikasi iPusnas di <i>Google Play Store</i> Menggunakan <i>Naive Bayes Classifier</i>	Penelitian tersebut menganalisis 700 ulasan pengguna iPusnas di <i>Google Play Store</i> meliputi algoritma <i>Naive Bayes</i> . Data diproses melalui tahap <i>preprocessing</i> dan dibagi dengan rasio 80:20. Hasil penelitian menunjukkan akurasi sebesar 58% dengan dominasi <i>sentimen</i> positif sebesar 75,1%, sehingga metode ini dapat dimanfaatkan untuk analisis <i>sentimen</i> dan evaluasi kualitas layanan.

Hak Cipta Dilindungi Undang-Undang

Universitas Islam Indragiri

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.

Plagiarisme juga dilarang dan dapat dikenakan sanksi. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



Tabel 2.1 Penelitian Terdahulu (Lanjutan)

NO	Nama Peneliti	Judul Penelitian	Pembahasan
(1)	(2)	(3)	(4)
10.	Elvin Alan Pratama, Aji Primajaya, Iqbal Maulana (2024)	Penerapan Algoritma <i>Support Vector Machine</i> Terhadap Analisis <i>Sentimen</i> Pada Ulasan Aplikasi LinkAja	Penelitian ini menerapkan algoritma <i>Support Vector Machine</i> (SVM) untuk menganalisis <i>sentimen</i> ulasan pengguna aplikasi LinkAja di Google <i>Play Store</i> . Dataset diperoleh melalui proses scraping sebanyak 2.000 ulasan. Proses penelitian meliputi tahapan Knowledge Discovery in Database (KDD) meliputi data <i>Selection</i> , <i>preprocessing</i> , transformation, data mining, evaluation, dan knowledge. Hasil penelitian menunjukkan model terbaik diperoleh pada pembagian data 70:30 dengan nilai <i>accuracy</i> 88%, <i>precision</i> 90%, <i>recall</i> 83%, dan <i>F1-score</i> 87%.

Penelitian oleh [10] yang berjudul Analisis *Sentimen* ulasan pengguna aplikasi Shopee pada Google *Play Store* menggunakan algoritma *Naive Bayes* berbasis Python. Penelitian ini bertujuan untuk mengeksplorasi opini pengguna terhadap aplikasi Shopee melalui Analisis *Sentimen* dari sekitar 1.300 ulasan yang dikumpulkan menggunakan teknik *web scraping*. Data yang diperoleh kemudian melalui tahapan *preprocessing*, meliputi pembersihan teks, *tokenisasi*, dan penghapusan *stopwords* untuk meningkatkan kualitas data sebelum dilakukan proses klasifikasi *sentimen*.





Berdasarkan hasil pengujian, algoritma Naive Bayes memperoleh tingkat akurasi sekitar 86%, yang menunjukkan kemampuan model dalam mengklasifikasikan sentimen ulasan pengguna dengan baik. Distribusi hasil klasifikasi memperlihatkan bahwa sentimen negatif lebih banyak dibandingkan sentimen positif. Kondisi tersebut mengindikasikan masih terdapat berbagai keluhan pengguna terhadap layanan aplikasi yang menjadi objek penelitian. Penggunaan data ulasan secara real-time memberikan gambaran yang lebih aktual mengenai persepsi pengguna, meskipun penelitian ini hanya berfokus pada satu platform dan belum membandingkan performa Naive Bayes dengan algoritma klasifikasi lainnya.

Penelitian selanjutnya oleh [11] bertujuan untuk menganalisis *sentimen* ulasan pengguna aplikasi *Jamsostek Mobile* (JMO) pada *Google Play Store* menggunakan algoritma *Naive Bayes*. Data diperoleh melalui teknik *web scraping* pada periode September hingga Oktober 2024, yang kemudian diproses melalui tahapan *Knowledge Discovery in Database* (KDD), meliputi seleksi data, *preprocessing* seperti *text cleaning*, penghapusan *stopwords*, normalisasi, dan *stemming*, serta transformasi data menggunakan metode *TF-IDF*. Selanjutnya, data diklasifikasikan ke dalam tiga kategori *sentimen*, yaitu positif, negatif, dan netral dengan pembagian data latih dan data uji sebesar 80:20 serta menggunakan teknik *cross-validation* untuk pengujian model.

Berdasarkan hasil klasifikasi, sentimen positif mendominasi dengan persentase 74,1%, sedangkan sentimen negatif sebesar 25,9%, sementara jumlah sentimen netral relatif sedikit. Model Naive Bayes menghasilkan akurasi sebesar 86,06% dengan nilai *precision*, *recall*, dan *F1-score* yang baik pada kategori

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.
2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



sentimen positif dan negatif. Namun, performa klasifikasi pada sentimen netral masih belum optimal karena jumlah data pada kategori tersebut jauh lebih sedikit dibandingkan kategori lainnya. Hasil tersebut menunjukkan bahwa algoritma Naive Bayes mampu memberikan performa yang baik dalam mengklasifikasikan sentimen pengguna, meskipun ketidakseimbangan distribusi data masih memengaruhi kemampuan model dalam mengenali kelas sentimen netral.

Penelitian selanjutnya oleh [12] bertujuan untuk menganalisis *sentimen* ulasan pengguna aplikasi Lita di Google *Play Store* menggunakan algoritma *Naive Bayes*. Data yang digunakan sebanyak 1000 ulasan yang diambil secara acak, kemudian diproses melalui tahapan *preprocessing* yang meliputi *cleaning*, *Tokenizing*, dan *stemming* untuk menghilangkan kata-kata yang tidak relevan. Selanjutnya, dilakukan pembobotan menggunakan metode *TF-IDF* untuk mengubah data teks menjadi representasi numerik sebelum dilakukan proses klasifikasi ke dalam dua kategori *sentimen*, yaitu positif dan negatif. Proses pengujian model dilakukan dengan pembagian data latih dan data uji dengan berbagai skenario, seperti 80:20, untuk mengukur performa model.

Berdasarkan hasil pengujian, algoritma Naive Bayes memperoleh tingkat akurasi sebesar 95%, dengan nilai *precision* dan *recall* yang juga mencapai 95%. Hasil klasifikasi menunjukkan bahwa mayoritas ulasan pengguna aplikasi Lita termasuk dalam kategori sentimen positif, yang mengindikasikan bahwa pengguna memberikan penilaian yang baik terhadap kualitas aplikasi. Meskipun demikian, proses klasifikasi hanya dilakukan pada dua kategori sentimen, yaitu positif dan negatif, sehingga hasil yang diperoleh masih dapat dikembangkan melalui

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

peraturan yang berlaku di Indonesia.

penggunaan kategori sentimen yang lebih beragam maupun penerapan dataset yang lebih luas.

Penelitian selanjutnya oleh [13] yang berjudul “Analisis *Sentimen* Pengguna terhadap Aplikasi Blu BCA di *Play Store* Menggunakan Algoritma *Naive Bayes*”. Penelitian ini bertujuan untuk menganalisis *sentimen* ulasan pengguna aplikasi Blu BCA serta mengidentifikasi persepsi pengguna terhadap aspek kinerja, keamanan, dan privasi aplikasi. Metode yang digunakan adalah Analisis *Sentimen* dengan algoritma *Naive Bayes*, dengan tahapan pengumpulan data melalui teknik *web scraping* dari Google *Play Store*, kemudian dilakukan *preprocessing* data yang meliputi data *cleaning*, *case folding*, *stopword removal*, *Tokenizing*, dan *stemming*, serta pembobotan menggunakan metode *TF-IDF*. Selanjutnya, data dibagi menjadi data latih dan data uji dengan proporsi 80:20 dan dievaluasi menggunakan *Confusion Matrix* untuk mengukur performa model.

Berdasarkan hasil pengujian, algoritma *Naive Bayes* memperoleh nilai akurasi sebesar 85,31% dengan nilai *precision*, *recall*, dan *F1-score* yang menunjukkan kinerja klasifikasi yang baik. Hasil analisis juga memperlihatkan bahwa sebagian besar ulasan pengguna memiliki sentimen positif. Meskipun demikian, masih ditemukan beberapa ulasan negatif yang berkaitan dengan performa aplikasi, kendala saat login, serta aspek keamanan. Temuan tersebut menunjukkan bahwa algoritma *Naive Bayes* mampu mengidentifikasi kecenderungan sentimen pengguna secara efektif dan menghasilkan informasi yang dapat dimanfaatkan sebagai bahan evaluasi dalam pengembangan layanan aplikasi perbankan digital.





Penelitian selanjutnya oleh [14] berjudul “Analisis *Sentimen* Ulasan Aplikasi DANA di Google *Play Store* Menggunakan Algoritma *Naïve Bayes*” yang bertujuan untuk menganalisis *Sentimen* pengguna serta mengetahui kecenderungan opini terhadap penggunaan aplikasi dompet digital DANA. Metode penelitian yang digunakan adalah *Knowledge Discovery from Data* (KDD) yang meliputi tahapan seleksi data, *preprocessing*, transformasi, data mining, dan evaluasi. Data dikumpulkan melalui teknik *web scraping* dari Google *Play Store* sebanyak 2.000 ulasan berbahasa Indonesia, kemudian dilakukan *preprocessing* yang mencakup *cleaning*, *case folding*, *tokenisasi*, *stopword removal*, dan *stemming*. Selanjutnya dilakukan pembobotan kata menggunakan *TF-IDF* serta proses klasifikasi menggunakan algoritma *Naïve Bayes*.

Berdasarkan hasil klasifikasi, sentimen positif menjadi kategori yang paling dominan dengan jumlah 1.476 ulasan, sedangkan sentimen negatif tercatat sebanyak 443 ulasan dan sentimen netral sebanyak 81 ulasan. Model klasifikasi menggunakan algoritma *Naïve Bayes* menghasilkan nilai akurasi 86%, dengan *precision* 86, *recall* 98%, dan *F1-score* 92%. Nilai tersebut menunjukkan bahwa model memiliki kemampuan yang baik dalam mengklasifikasikan sentimen pengguna. Ulasan yang termasuk dalam kategori negatif umumnya berkaitan dengan kendala saat login, kegagalan transaksi, serta permasalahan pada proses pengisian saldo, sehingga aspek-aspek tersebut menjadi masukan yang dapat dipertimbangkan dalam pengembangan dan peningkatan kualitas layanan aplikasi.

Penelitian selanjutnya oleh [15] berjudul “Analisis *Sentimen* Pengguna Aplikasi Agoda di Google *Play Store* Menggunakan Algoritma *Naïve Bayes*” yang bertujuan untuk menganalisis *sentimen* pengguna serta mengevaluasi

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

peraturan yang berlaku di Indonesia.



kualitas layanan aplikasi Agoda berdasarkan ulasan pengguna. Metode penelitian yang digunakan meliputi tahapan pengumpulan data, *preprocessing*, pelabelan data menggunakan pendekatan lexicon-based, pembobotan *TF-IDF*, pembagian data latih dan data uji, serta proses klasifikasi menggunakan algoritma *Naïve Bayes*. Data dikumpulkan melalui teknik *web scraping* dari Google Play Store sebanyak 7.885 ulasan berbahasa Indonesia, kemudian dilakukan *preprocessing* yang mencakup *cleaning*, *case folding*, *Tokenizing*, *normalization*, *stopword removal*, dan *stemming* untuk meningkatkan kualitas data sebelum proses klasifikasi dilakukan.

Hasil penelitian menunjukkan bahwa *sentimen* negatif lebih mendominasi dibandingkan *sentimen* positif, dengan jumlah 4.733 ulasan negatif dan 2.962 ulasan positif. Model yang dibangun mampu mencapai akurasi terbaik sebesar 91,9% pada skenario pembagian data tertentu, yang menunjukkan bahwa algoritma *Naïve Bayes* efektif dalam mengklasifikasikan *sentimen* pengguna. Selain itu, ditemukan bahwa keluhan pengguna umumnya berkaitan dengan masalah pembatalan, pengembalian dana, proses pemesanan, serta ketidakpuasan terhadap harga dan layanan. Penelitian ini menegaskan bahwa Analisis *Sentimen* dapat memberikan wawasan penting bagi pengembang dalam meningkatkan kualitas layanan aplikasi serta dapat dijadikan acuan untuk penelitian selanjutnya dalam pengembangan metode analisis *sentimen*.

Penelitian selanjutnya oleh [16] berjudul “Analisis *Sentimen* pada Ulasan Aplikasi Wondr di *Play Store* dengan Metode *Naïve Bayes*” yang bertujuan untuk menganalisis persepsi serta opini pengguna terhadap aplikasi mobile banking Wondr sebagai layanan digital dari Bank Negara Indonesia (BNI). Metode

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

peraturan yang berlaku di Indonesia.



penelitian yang digunakan adalah kuantitatif dengan pendekatan Analisis Sentimen menggunakan algoritma *Naive Bayes Classifier* yang diimplementasikan melalui bahasa pemrograman Python. Data penelitian dikumpulkan melalui teknik *web scraping* dari Google Play Store menggunakan Google Colab sebanyak 27.124 ulasan pengguna, kemudian dilakukan tahapan *preprocessing* yang meliputi *cleaning*, *tokenization*, *case folding*, *stopword removal*, filtering token by length, dan *stemming*. Selanjutnya data diberi label ke dalam tiga kategori sentimen (positif, negatif, dan netral), dilakukan pembobotan menggunakan *TF-IDF*, serta evaluasi model menggunakan *Confusion Matrix* dengan metrik akurasi, *precision*, *recall*, dan *F1-score*.

Berdasarkan hasil klasifikasi, sentimen positif menjadi kategori yang paling dominan dengan persentase 52,9%, diikuti sentimen negatif sebesar 39,9%, sedangkan sentimen netral hanya mencapai 7,2%. Model Naive Bayes menghasilkan tingkat akurasi sebesar 82%, yang menunjukkan kemampuan yang cukup baik dalam mengelompokkan sentimen pengguna berdasarkan ulasan di Google Play Store. Meskipun demikian, kemampuan model dalam mengenali sentimen netral masih belum optimal karena jumlah data pada kategori tersebut relatif lebih sedikit dibandingkan kategori lainnya. Hasil analisis juga menunjukkan bahwa sebagian besar pengguna memberikan tanggapan positif terhadap aplikasi Wondr, namun masih terdapat beberapa ulasan yang menyoroti kendala pada sistem, pengalaman penggunaan, serta fitur aplikasi yang memerlukan penyempurnaan.

Penelitian selanjutnya oleh [17] berjudul “Analisis Sentimen Pengguna Aplikasi Bank Syariah Indonesia dengan Menggunakan Algoritma *Support Vector*

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.
2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber. Plagiarisme juga dilarang dan dapat dikenakan sanksi.
3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.



Machine (SVM)” yang bertujuan untuk menganalisis *sentimen* serta tingkat kepuasan pengguna terhadap aplikasi BSI Mobile di Google *Play Store*. Metode penelitian yang digunakan adalah kuantitatif dengan pendekatan Analisis *Sentimen* menggunakan algoritma *Support Vector Machine (SVM)*. Data penelitian dikumpulkan melalui teknik *web scraping* dari Google *Play Store* pada periode 1 Februari 2021 hingga 10 Agustus 2022, kemudian dilakukan pelabelan secara manual oleh lima ahli bahasa sehingga diperoleh sebanyak 13.568 data ulasan. Data tersebut selanjutnya melalui tahapan *preprocessing* yang meliputi *case folding*, *filtering*, *stopword removal*, *tokenization*, dan *stemming*, serta dibagi menjadi data training sebesar 70% dan data testing sebesar 30%.

Berdasarkan hasil pengujian, model *Support Vector Machine (SVM)* memperoleh nilai akurasi sebesar 85,87%, dengan *precision* sebesar 87,309% dan *recall* sebesar 86,958%. Hasil klasifikasi menunjukkan bahwa *sentimen* positif lebih banyak dibandingkan *sentimen* negatif, yang mengindikasikan bahwa mayoritas pengguna memberikan penilaian yang baik terhadap aplikasi BSI Mobile. Meskipun demikian, masih ditemukan sejumlah ulasan yang mengungkapkan adanya gangguan sistem serta penurunan performa aplikasi. Temuan tersebut menunjukkan bahwa metode SVM memiliki kemampuan yang baik dalam mengklasifikasikan *sentimen* pengguna sekaligus memberikan informasi mengenai aspek layanan yang masih perlu ditingkatkan.

Penelitian selanjutnya oleh [18] dengan judul “Analisis *Sentimen* pada Ulasan Aplikasi iPusnas di Google *Play Store* Menggunakan *Naive Bayes Classifier*.” Penelitian ini bertujuan untuk menganalisis dan mengklasifikasikan *sentimen* ulasan pengguna aplikasi iPusnas yang diperoleh dari Google *Play Store*

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan tanpa izin untuk kepentingan komersial atau pelanggaran hak cipta dapat dikenakan sanksi sesuai dengan UU Hak Cipta di Indonesia.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

peraturan yang berlaku di Indonesia.



menggunakan algoritma *Naive Bayes*. Data yang digunakan sebanyak 700 ulasan pengguna yang telah melalui tahapan *preprocessing*, seperti *case folding*, *tokenization*, dan *stemming*, kemudian dibagi menjadi data latih dan data uji dengan rasio 80:20 .

Model yang dibangun selanjutnya dievaluasi menggunakan *Confusion Matrix* untuk mengukur kinerja klasifikasi. Hasil penelitian menunjukkan bahwa sebagian besar ulasan pengguna memiliki *sentimen* positif sebesar 75,1%, yang mengindikasikan tingkat kepuasan pengguna yang cukup tinggi. Selain itu, algoritma *Naive Bayes* menghasilkan akurasi sebesar 58% dengan nilai *precision* 69%, *recall* 81%, dan *F1-score* 75%. Penelitian ini menunjukkan bahwa metode *Naive Bayes* cukup efektif dalam mengklasifikasikan *sentimen* ulasan pengguna serta dapat dimanfaatkan sebagai bahan evaluasi dalam meningkatkan kualitas layanan aplikasi iPusnas

Penelitian selanjutnya oleh [19] berjudul “Penerapan Algoritma *Support Vector Machine* terhadap Analisis *Sentimen* pada Ulasan Aplikasi LinkAja” yang bertujuan untuk menganalisis *sentimen* pengguna serta menginterpretasikan hasilnya sebagai rekomendasi dalam meningkatkan kualitas layanan aplikasi LinkAja. Metode penelitian yang digunakan adalah Knowledge Discovery in Database (KDD) dalam text mining, yang meliputi tahapan data *Selection*, *preprocessing*, transformation, data mining, evaluation, dan knowledge. Data dikumpulkan melalui teknik *web scraping* dari Google Play Store sebanyak 2.000 ulasan pengguna, kemudian dilakukan *preprocessing* yang mencakup *cleaning*, *case folding*, *Tokenizing*, normalization, *stopword removal*, dan *stemming*. Selanjutnya, dilakukan pelabelan data menggunakan *lexicon InSet*, serta

1. Dilarang memperbanyak atau mendistribusikan dokumen ini untuk tujuan komersial tanpa izin tertulis dari penulis atau pihak berwenang.

2. Penggunaan tanpa izin untuk kepentingan akademik, penelitian, dan pendidikan diperbolehkan dengan mencantumkan sumber.

3. Universitas hanya berhak menyimpan dan mendistribusikan dokumen ini di repositori akademik, tanpa mengalihkan hak cipta penulis, sesuai dengan peraturan yang berlaku di Indonesia.

peraturan yang berlaku di Indonesia.

pembobotan menggunakan *TF-IDF* sebelum proses klasifikasi menggunakan algoritma *Support Vector Machine* (SVM).

Berdasarkan hasil klasifikasi, sentimen negatif lebih dominan dengan jumlah 1.056 ulasan, sedangkan sentimen positif tercatat sebanyak 944 ulasan. Performa terbaik model *Support Vector Machine* (SVM) diperoleh pada skenario pembagian data 70:30, dengan nilai akurasi 88%, precision 90%, recall 83%, dan F1-score 87%. Analisis terhadap ulasan pengguna juga menunjukkan bahwa sebagian besar keluhan berkaitan dengan kurang optimalnya layanan pelanggan, aplikasi yang sering mengalami crash pada perangkat dengan spesifikasi rendah, kendala koneksi, serta tampilan antarmuka yang dinilai kurang nyaman digunakan. Hasil tersebut menunjukkan bahwa metode SVM mampu mengklasifikasikan sentimen pengguna dengan baik sekaligus mengidentifikasi aspek-aspek layanan yang masih memerlukan perbaikan.

2.6 Rangkuman

Penelitian ini bertujuan untuk menganalisis *sentimen* pengguna terhadap aplikasi DOLA di Google *Play Store* menggunakan algoritma *Naive Bayes*. Data penelitian diperoleh melalui teknik *web scraping* sebanyak 5.000 ulasan pengguna, kemudian dilakukan proses data *Selection* dan data labeling menjadi tiga kategori *sentimen*, yaitu positif, netral, dan negatif, berdasarkan nilai rating pengguna.

Sebelum proses klasifikasi, data melalui tahapan *preprocessing* yang meliputi *case folding*, *cleaning*, *tokenization*, *stopword removal*, dan *stemming* untuk membersihkan serta menormalkan teks. Selanjutnya, data diubah ke dalam bentuk numerik menggunakan metode *TF-IDF*, kemudian dilakukan *oversampling* untuk



menyeimbangkan distribusi kelas sehingga model dapat belajar dengan lebih optimal.

Proses klasifikasi dilakukan menggunakan algoritma *Naive Bayes* dengan pembagian data 80% sebagai data latih dan 20% sebagai data uji, serta evaluasi menggunakan *k-fold cross-validation* untuk memperoleh hasil yang lebih stabil. Evaluasi model dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *Confusion Matrix* untuk mengukur kinerja klasifikasi.

Hasil penelitian menunjukkan bahwa model mampu mencapai akurasi sebesar 82,37%, yang menandakan bahwa algoritma *Naive Bayes* cukup efektif dalam mengklasifikasikan *sentimen* pengguna aplikasi DOLA. Nilai evaluasi juga menunjukkan kinerja yang baik, dengan *weighted average precision* sebesar 0,83, *recall* sebesar 0,82, dan *F1-score* sebesar 0,82. Selain itu, hasil *Confusion Matrix* memperlihatkan bahwa sebagian besar data berhasil diprediksi dengan benar, meskipun masih terdapat beberapa kesalahan klasifikasi antar kelas *sentimen*.

Secara keseluruhan, penelitian ini membuktikan bahwa kombinasi *preprocessing*, *TF-IDF*, dan algoritma *Naive Bayes* mampu memberikan hasil klasifikasi yang cukup baik dalam Analisis *Sentimen* ulasan pengguna aplikasi DOLA. Hasil penelitian ini diharapkan dapat menjadi bahan evaluasi bagi pengembang untuk meningkatkan kualitas layanan aplikasi serta menjadi referensi bagi penelitian selanjutnya di bidang analisis *sentimen*.

