



BAB 1

PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi informasi telah mengubah berbagai aspek kehidupan, termasuk cara masyarakat memperoleh informasi, berkomunikasi, dan memanfaatkan layanan digital dalam aktivitas sehari-hari. Meningkatnya akses internet turut mempercepat penggunaan berbagai aplikasi berbasis perangkat seluler yang memudahkan pengguna memperoleh layanan secara praktis. Seiring dengan meningkatnya penggunaan aplikasi digital, jumlah data yang dihasilkan juga semakin besar dan beragam. Data-data tersebut dapat memberikan informasi yang bermanfaat apabila diolah dan dianalisis dengan baik. Salah satu platform yang menyediakan berbagai aplikasi digital yaitu *Google Play Store*, yang memungkinkan pengguna menggunakan aplikasi, pengguna dapat memberikan penilaian dalam bentuk rating serta ulasan sebagai bentuk umpan balik terhadap kualitas layanan. Informasi yang terdapat pada ulasan pengguna dapat dimanfaatkan untuk memahami pengalaman penggunaan aplikasi sekaligus mengidentifikasi tingkat kepuasan maupun permasalahan yang dirasakan pengguna [1].

Seiring dengan meningkatnya penggunaan aplikasi mobile, jumlah ulasan diberikan oleh pengguna pada platform seperti *Google Play Store* juga semakin banyak. Ulasan tersebut biasanya berisi berbagai pendapat, pengalaman, kritik, maupun saran terhadap kinerja dan kualitas suatu aplikasi. Informasi yang terkandung dalam ulasan pengguna tersebut penting bagi pengembang aplikasi untuk mengetahui tingkat kepuasan pengguna serta mengidentifikasi berbagai



permasalahan yang mungkin terjadi pada aplikasi yang mereka kembangkan. Namun, banyaknya jumlah ulasan yang tersedia seringkali menyulitkan proses analisis apabila dilakukan secara manual. Oleh karena itu, diperlukan suatu metode yakni mampu mengolah dan menganalisis data ulasan pengguna secara otomatis sehingga bisa menghasilkan informasi yang lebih efektif dan efisien[2].

Banyaknya ulasan yang dihasilkan pengguna membuat proses analisis secara manual menjadi kurang efisien. Oleh karena itu, diperlukan metode komputasi yang mampu mengelompokkan opini pengguna secara otomatis berdasarkan isi ulasan. Salah satu teknik yakni dapat digunakan adalah analisis *sentimen*, yaitu metode dalam pengolahan teks yang bertujuan untuk mengekstrak, mengidentifikasi, maupun mengklasifikasikan opini atau pendapat yang terdapat dalam suatu teks ke dalam kategori *sentimen* seperti positif, negatif, dan netral. Analisis *Sentimen* banyak digunakan dalam berbagai bidang seperti e-commerce, media sosial, maupun layanan digital untuk mengetahui persepsi pengguna terhadap suatu produk atau layanan. Dengan memanfaatkan teknik analisis *sentimen*, data ulasan pengguna dengan berjumlah besar dapat dianalisis secara lebih sistematis sehingga menghasilkan informasi yakni dapat membantu pengembang dalam mengevaluasi dan meningkatkan kualitas aplikasi[3].

Dalam proses analisis *sentimen*, diperlukan suatu metode klasifikasi mampu mengelompokkan data teks ke dalam kategori *sentimen* yang telah ditentukan, seperti positif, negatif, maupun netral. Salah satu algoritma yang banyak digunakan dalam penelitian *Analisis Sentimen* adalah *Naïve Bayes*, yaitu metode klasifikasi berbasis probabilitas yaitu mengacu pada Teorema Bayes untuk menentukan kemungkinan suatu data termasuk ke dalam kelas tertentu. Algoritma

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.



ini banyak diterapkan karena memiliki keunggulan dalam kesederhanaan perhitungan, efisiensi dalam pengolahan data teks, serta kemampuan menghasilkan performa yang cukup baik meskipun menggunakan dataset yang relatif sederhana. Selain itu, *Naïve Bayes* juga efektif digunakan pada data berukuran besar dan tidak terstruktur seperti data media social. Dengan memanfaatkan algoritma ini, data berupa ulasan atau opini pengguna dapat diklasifikasikan secara otomatis sebagai kategori *sentimen* yang sesuai, sehingga dapat membantu dalam memahami kecenderungan opini publik pada suatu topik, produk, atau layanan secara lebih cepat, sistematis, dan mendukung pengambilan keputusan berbasis data[4].

Berbagai penelitian sebelumnya menunjukkan bahwa Analisis *Sentimen* berdasarkan ulasan pengguna pada aplikasi mobile di Google *Play Store* dapat memberikan informasi penting mengenai tingkat kepuasan pengguna serta berbagai permasalahan yang terdapat pada aplikasi tersebut. Ulasan pengguna sering kali memuat berbagai keluhan seperti kesulitan untuk mengakses aplikasi, gangguan teknis, maupun keterbatasan fitur yang tersedia. Oleh karena itu, Analisis *Sentimen* menjadi salah satu pendekatan sehingga efektif untuk memahami persepsi pengguna terhadap suatu aplikasi. Dengan memanfaatkan metode klasifikasi seperti *Naive Bayes*, data ulasan yaitu diperoleh dari Google *Play Store* dapat dianalisis untuk mengetahui kecenderungan *sentimen* pengguna serta memberikan rekomendasi bagi pengembang untuk meningkatkan kualitas layanan aplikasi[5].

Berdasarkan permasalahan tersebut, penelitian ini dilakukan dengan tujuan sebagai menganalisis *sentimen* pengguna terhadap aplikasi DOLA berdasarkan

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. 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.



ulasan yang terdapat pada platform *Google Play Store* dengan menggunakan algoritma *Naive Bayes*. Melalui penelitian ini, ulasan pengguna akan diproses dan diklasifikasikan secara otomatis terhadap kategori *sentimen* positif, negatif, dan netral sehingga dapat memberikan gambaran mengenai persepsi serta tingkat kepuasan pengguna terhadap aplikasi DOLA. Dengan adanya Analisis *Sentimen* ini, diharapkan dapat diketahui kecenderungan opini pengguna terhadap kinerja, fitur, maupun kualitas layanan yang diberikan oleh aplikasi tersebut. Hasil dari penelitian tersebut diharapkan dapat memberikan manfaat secara teoritis maupun praktis. Secara teoritis, penelitian ini diharapkan dapat menambah referensi dan wawasan dalam bidang text mining khususnya pada penerapan metode *Naive Bayes* sebagai Analisis *Sentimen* terhadap data ulasan pengguna aplikasi. Sementara secara praktis, hasil penelitian ini diharapkan dapat menjadi bahan evaluasi dan masukan bagi pihak pengembang aplikasi DOLA dalam meningkatkan kualitas layanan, memperbaiki kekurangan aplikasi, serta meningkatkan kepuasan pengguna di masa yang akan datang.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang sebelumnya, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Belum adanya analisis yang mendalam mengenai *sentimen* pengguna terhadap aplikasi Dola berdasarkan ulasan di *Google Play Store*.
2. Belum adanya penerapan algoritma *Naive Bayes* secara khusus untuk mengklasifikasikan *sentimen* pada ulasan pengguna aplikasi Dola.
3. Belum adanya pengujian akurasi algoritma *Naive Bayes* dalam menganalisis dan mengidentifikasi *sentimen* pengguna terhadap aplikasi Dola di *Play Store*.



1.3 Batasan Masalah

Agar penelitian ini lebih terarah dan tidak menyimpang dari tujuan, maka ditetapkan batasan masalah sebagai berikut:

1. Penelitian ini hanya berfokus pada Analisis *Sentimen* pengguna terhadap aplikasi Dola yang diambil dari ulasan pada platform *Google Play Store*.
2. Data yang digunakan terbatas pada ulasan berbahasa Indonesia dan diperoleh dalam rentang waktu tertentu sesuai kebutuhan penelitian.
3. Metode klasifikasi *sentimen* yang digunakan dalam penelitian ini dibatasi pada penerapan algoritma *Naive Bayes*, tanpa membandingkannya dengan algoritma lain.
4. Kategori *sentimen* ini digunakan dalam penelitian yang dibatasi menjadi tiga kelas, yaitu positif, negatif, dan netral.
5. Evaluasi performa model hanya dilakukan menggunakan metrik akurasi, presisi, *recall*, dan *F1-score* untuk menilai tingkat keberhasilan algoritma *Naive Bayes* dalam klasifikasi *sentimen*.

1.4 Tujuan Penelitian

Adapun tujuan penelitian ini untuk menggambarkan hasil yang telah direncanakan. Antara lain yakni:

1. Mengetahui dan menganalisis *sentimen* pengguna terhadap aplikasi Dola berdasarkan ulasan yang terdapat di *Google Play Store*.
2. Mengklasifikasikan ulasan pengguna terhadap kategori *sentimen* positif, negatif, dan netral menggunakan algoritma *Naive Bayes* meliputi tahapan *preprocessing*, pelabelan data, dan pembobotan *TF-IDF*.



3. Mengevaluasi kinerja model *Naive Bayes* terhadap mengklasifikasikan *sentimen* pengguna menggunakan metrik akurasi, presisi, *recall*, dan *F1-score* untuk mengetahui tingkat efektivitas model.

1.5 Manfaat Penelitian

Adapun manfaat dari penelitian ini:

1. Memberikan gambaran mengenai *sentimen* pengguna aplikasi Dola berdasarkan ulasan di *Google Play Store* sehingga dapat digunakan sebagai acuan dalam memahami persepsi pengguna.
2. Menyediakan hasil klasifikasi *sentimen* yang dapat membantu pengembangan aplikasi dalam mengevaluasi kualitas layanan dan meningkatkan fitur yang dianggap kurang oleh pengguna.
3. Menjadi referensi terhadap peneliti atau pihak lain yang ingin melakukan penelitian terkait analisis *sentimen*, khususnya yang menggunakan algoritma *Naive Bayes* dan metode *TF-IDF* sebagai pendekatan pengolahan teks.

1.6 Sistematika Penelitian

Sistematika penulisan yang digunakan pada penelitian ini adalah sebagai berikut:

BAB I : PENDAHULUAN

Pada bab I ini terdapat latar belakang masalah yang menjadi latar belakang penelitian ini dilakukan, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

BAB II : TINJAUAN LITERATUR

Bab II ini terdapat 10 jurnal penelitian terdahulu yang berkaitan dengan Analisis *Sentimen* dan algoritma *Naive Bayes* untuk dibandingkan dengan penelitian yang dilakukan. Pada bagian akhir disajikan rangkuman hasil tinjauan penelitian terdahulu sebagai dasar penguatan teori dan identifikasi celah penelitian.

BAB III : METODE PENELITIAN

Bab ini menguraikan metode yang diterapkan dalam penelitian menggunakan pendekatan kuantitatif deskriptif dengan algoritma *Naive Bayes* untuk melakukan klasifikasi sentimen. Tahapan penelitian meliputi penyusunan kerangka penelitian, identifikasi permasalahan, pengambilan data ulasan aplikasi DOLA melalui teknik web scraping pada Google Play Store, seleksi data, preprocessing teks, pembobotan menggunakan TF-IDF, pelabelan sentimen, pembentukan model *Naive Bayes*, serta evaluasi kinerja model berdasarkan nilai akurasi, presisi, recall, dan F1-score.

BAB IV : HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil penelitian dan pembahasan proses Analisis *Sentimen* pengguna aplikasi DOLA di Google Play Store menggunakan algoritma *Naive Bayes*. Pada bab ini dijelaskan tahapan penelitian dimulai dari proses pengumpulan data ulasan menggunakan teknik *web scraping*, data *Selection*, pelabelan *sentimen*, serta *preprocessing* data yang meliputi *case folding*, *cleaning*, *tokenization*, *stopword removal*, dan *stemming*. Selanjutnya, data teks ditransformasikan menerapkan metode *TF-IDF* sebagai bentuk numerik untuk kemudian dilakukan proses klasifikasi menggunakan algoritma *Naive Bayes*. Setelah itu, dilakukan pembagian data sebagai data training dan data testing





dengan beberapa skenario pembagian data untuk mengetahui pengaruh perbedaan proporsi data terhadap kinerja model. Evaluasi model dilakukan menggunakan *Confusion Matrix* dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, serta pengujian menggunakan metode k-fold cross validation untuk memperoleh hasil evaluasi yang lebih stabil. Pada tahap akhir, dilakukan visualisasi *Word Cloud* guna menampilkan kata-kata yang paling sering muncul pada setiap kategori *sentimen* sehingga dapat memberikan gambaran mengenai kecenderungan opini pengguna terhadap aplikasi DOLA.

BAB V : KESIMPULAN DAN SARAN

Pada bab ini merupakan bagian akhir dari penelitian yang berisi kesimpulan dan saran berdasarkan hasil Analisis *Sentimen* pengguna aplikasi DOLA menggunakan algoritma *Naive Bayes*. Kesimpulan berisi rangkuman hasil penelitian, termasuk performa model klasifikasi yang diperoleh berdasarkan nilai *accuracy*, *precision*, *recall*, dan *F1-score* serta gambaran kecenderungan *sentimen* pengguna terhadap aplikasi DOLA. Selanjutnya, bagian saran berisi rekomendasi ini dapat dipakai sebagai bahan pengembangan penelitian berikutnya, seperti penggunaan jumlah data yang lebih besar, penerapan metode klasifikasi lain, maupun peningkatan tahapan *preprocessing* agar menghasilkan performa Analisis *Sentimen* yang lebih optimal.

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. 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.