

BAB I

PENDAHULUAN

1.1 Latar Belakang

Penggunaan internet di Indonesia terus meningkat pada tiga tahun terakhir mulai tahun 2021 hingga 2023. Hal tersebut berdasarkan laporan data dari Data Reportal. Diawali pada Januari 2021 dengan jumlah pengguna internet yaitu 203 juta pengguna. Kemudian pada Januari 2022 dengan jumlah pengguna internet yaitu 205 juta pengguna. Selanjutnya pada Januari 2023 sebesar 212,9 juta jumlah pengguna internet di Indonesia (Kemp, 2023). Data tersebut menunjukkan bahwa internet dapat membantu aktivitas kegiatan manusia sehari-hari sehingga jumlah pengguna internet terus meningkat. Salah satu contoh penggunaan internet yaitu sebagai media untuk berkomunikasi antar individu manusia dengan memanfaatkan media sosial. Manfaat media sosial tidak hanya sebagai media untuk berkomunikasi, media sosial juga dapat digunakan sebagai media untuk pencarian pertemanan maupun pasangan. Salah satu platform yang memiliki fungsi tersebut adalah aplikasi bernama Bumble.

Aplikasi Bumble merupakan aplikasi yang didirikan oleh Whitney Wolfe Herd. Aplikasi tersebut telah berkembang jauh melampaui aplikasi kencan menjadi platform jaringan, memungkinkan orang dari semua jenis kelamin untuk membuat koneksi yang diberdayakan di semua bidang kehidupan mereka. Pengguna yang sedang mencari hubungan romantis dapat menggunakan fitur Bumble Date, sedangkan pengguna yang ingin menjalin pertemanan baru di Bumble BFF. Bumble telah menjadi aplikasi nomor satu pada kategori aplikasi kencan dan terlaris di *Google Play Store* pada Februari 2023. Aplikasi ini juga telah diunduh sebanyak 50 juta kali oleh pengguna dan memiliki rating 4.6 dari skala 1 sampai dengan 5. Selain itu, aplikasi Bumble juga telah menjadi salah satu aplikasi yang memenangi *App Store Awards 2021* pada kategori “*Trend of 2021: Apps That Brought Us Together*” di *App Store*. Pencapaian aplikasi Bumble menunjukkan bahwa aplikasi ini sangat berguna dalam kehidupan sosial penggunanya.

Aplikasi Bumble tersedia pada *Google Play Store* dan *App Store*. Pada platform *Google Play Store*, pengguna aplikasi dapat memberikan ulasan terhadap aplikasi tersebut serta memberikan skala penilaian terhadap aplikasi. Ulasan yang diberikan oleh pengguna beragam. Ulasan tersebut berisi kritik, saran, maupun pujian mengenai aplikasi tersebut. Ulasan tersebut juga digunakan oleh pengembang aplikasi sebagai bahan evaluasi untuk mengembangkan aplikasi sesuai pengalaman pengguna aplikasi tersebut. Ulasan-ulasan pengguna aplikasi Bumble akan dijadikan sebagai bahan penelitian mengenai topik *Natural Language Processing*(NLP).

Salah satu penelitian NLP yang populer adalah penelitian *sentiment analysis*. *Sentiment Analysis* merupakan proses komputasi untuk melakukan klasifikasi terhadap data tekstual sesuai dengan sentimen yang dimiliki (Sumiah & Mirantika, 2020). Terdapat berbagai macam metode *supervised machine learning* untuk melakukan *Sentiment Analysis*. Dalam *supervised machine learning*, algoritma komputer dilatih dengan *input* data yang telah diberi label khusus sehingga menghasilkan *output* tertentu. Metode *supervised machine learning* yang diimplementasikan pada penelitian ini adalah algoritma *K-Nearest Neighbor*(KNN). Algoritma *K-Nearest Neighbor* adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada.

Peneliti memilih algoritma KNN sebagai metode klasifikasi yang digunakan karena berdasarkan penelitian sebelumnya analisis sentimen menggunakan algoritma KNN memiliki hasil akurasi yang cukup baik. Pada penelitian sebelumnya yang dilakukan oleh Sumiah & Mirantika(2020) dan Alusius(2021) menunjukkan algoritma KNN memiliki nilai akurasi lebih tinggi dibandingkan metode *supervised machine learning* lainnya yang digunakan pada penelitian tersebut. Selain itu, untuk memaksimalkan penelitian juga dilakukan implementasi sebuah metode *word embedding*.

Word embedding merupakan sebuah fungsi parameter yang memetakan setiap kata ke dalam vektor berdimensi tinggi. Salah satu macam *word embedding* adalah *Word2vec*. *Word2vec* merepresentasikan kata ke dalam bentuk vektor

dengan ukuran n-dimensi tertentu. Peneliti menggunakan metode *word embedding* *Word2vec* pada penelitian ini berdasarkan hasil penelitian yang dilakukan oleh Rusyda(2021) yang menunjukkan peningkatan nilai *accuracy* dengan menggunakan metode *Word2vec* dibandingkan dengan metode *word embedding* lainnya.

Dari perihal tersebut, hasil penelitian *sentiment analysis* ulasan-ulasan pengguna pada aplikasi *online dating* Bumble dengan metode *word embedding* *Word2vec* dan algoritma *K-Nearest Neighbor* dapat menjadi bahan pertimbangan evaluasi dan pengembangan aplikasi selanjutnya sehingga aplikasi dapat digunakan lebih nyaman oleh pengguna.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka dapat dibentuk beberapa poin rumusan masalah yaitu:

1. Bagaimana mengimplementasikan algoritma *K-Nearest Neighbor* dan metode *word embedding* *Word2vec* untuk menganalisis sentimen ulasan pengguna tentang aplikasi *online dating* Bumble di *Google Play Store*?
2. Bagaimana pengaruh nilai k pada penggunaan algoritma *K-Nearest Neighbor* terhadap tingkat *accuracy*, *precision*, *recall*, dan *f1 score* pada analisis sentimen ulasan pengguna aplikasi *online dating* Bumble di *Google Play Store*?
3. Bagaimana pengaruh ukuran vektor dan metode pembuat vektor kata *Skip-gram* dan *Continues Bag Of Words*(CBOW) terhadap tingkat *accuracy*, *precision*, *recall*, dan *f1 score* pada analisis sentimen ulasan pengguna aplikasi *online dating* Bumble di *Google Play Store*?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang ada, maka tujuan dari penelitian ini adalah:

1. Mengimplementasikan algoritma *K-Nearest Neighbor* dan metode *word embedding Word2vec* untuk menganalisis sentimen ulasan-ulasan pengguna tentang aplikasi *online dating* Bumble di *Google Play Store*.
2. Mengetahui pengaruh nilai k pada penggunaan algoritma *K-Nearest Neighbor* terhadap tingkat *accuracy*, *precision*, *recall*, dan *f1 score* pada analisis sentimen ulasan pengguna aplikasi *online dating* Bumble di *Google Play Store*.
3. Mengetahui pengaruh ukuran vektor dan metode pembuat vektor kata *Skip-gram* dan *Continues Bag Of Words(CBOW)* terhadap tingkat *accuracy*, *precision*, *recall*, dan *f1 score* pada analisis sentimen ulasan pengguna aplikasi *online dating* Bumble di *Google Play Store*.

1.4 Manfaat Penelitian

Berikut merupakan manfaat yang ingin dicapai dari penelitian ini:

1. Bagi perusahaan yaitu sebagai bahan evaluasi untuk mengembangkan aplikasi *online dating* Bumble menjadi aplikasi yang lebih baik lagi selanjutnya.
2. Bagi mahasiswa yaitu mahasiswa dapat mengimplementasikan ilmu-ilmu yang telah diperoleh semasa kuliah serta dapat mempelajari dan mengimplementasikan mengenai penggunaan algoritma *K-Nearest Neighbor* dan metode *word embedding Word2vec* untuk menganalisis sentimen ulasan-ulasan pengguna tentang aplikasi *online dating* Bumble di *Google Play Store*.

1.5 Batasan Masalah

Berikut batasan masalah yang ditetapkan pada penelitian ini yaitu:

1. Penelitian hanya terbatas pada pengujian laboratorium.
2. Penelitian ini menggunakan algoritma *K-Nearest Neighbor*(KNN) dan metode *word embedding Word2vec*.
3. Pengimplementasian algoritma menggunakan bahasa pemrograman Python.
4. Penelitian ini menggunakan dataset sejumlah 1200 buah data dengan menggunakan data ulasan pengguna aplikasi *online dating* Bumble di *Google Play Store*.
5. Data ulasan yang digunakan pada penelitian ini yaitu data ulasan dengan bahasa Indonesia