

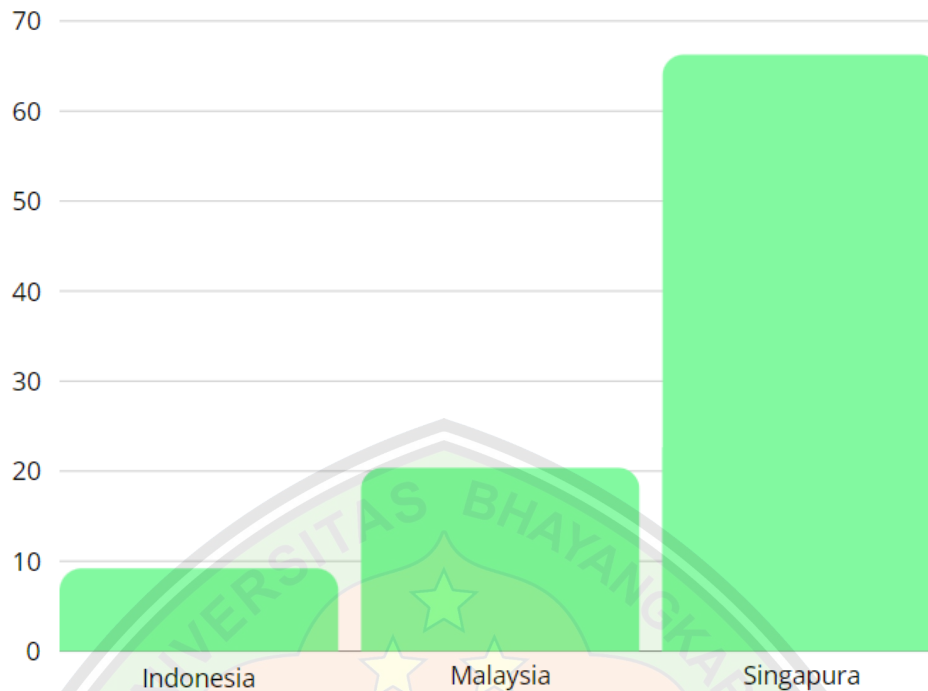
BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Saat ini adalah era yang dipenuhi oleh teknologi yang canggih dan sangat berkembang yang memungkinkan kita dapat mencari relasi dan informasi dengan melalui media sosial. Terdapat banyak media sosial namun yang umum digunakan seperti LinkedIn, Facebook, Twitter, dan Instagram. Media sosial merupakan platform daring (*online*) di mana pengguna dapat berpartisipasi dan menciptakan konten termasuk blog, jejaring sosial, wiki, forum, dan dunia maya dengan sangat mudah [1].

LinkedIn adalah *platform* media sosial yang berfokus pada bisnis dan pekerjaan melalui website dan aplikasi *mobile*. *Platform* linkedin ini utamanya digunakan oleh jaringan profesional dan pengembangan karir memungkinkan pencari kerja untuk menampilkan *curriculum vitae* mereka, serta perusahaan atau pemberi kerja menampilkan lowongan pekerjaan [2].

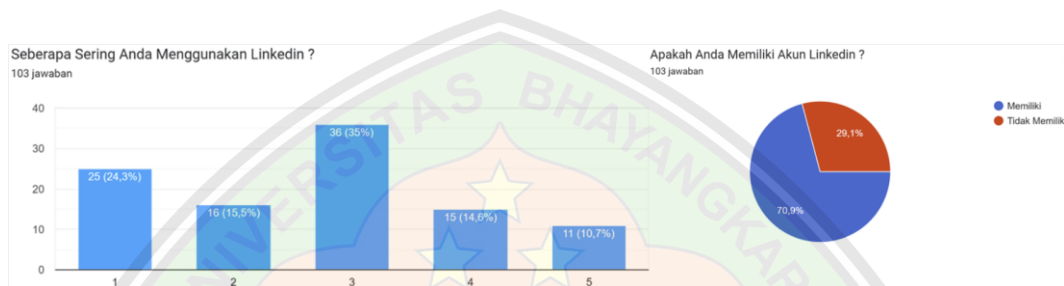


Gambar 1. 1 Rasio pengguna LinkedIn di ketiga negara

Di Indonesia sebanyak 26,7 juta pengguna *platform LinkedIn*. Jika dibandingkan dengan penduduk Indonesia sebanyak 277,5 juta jiwa, maka pengguna *platform LinkedIn* memiliki rasio sebesar 9.2%. Persentase ini sangatlah kecil jika dibandingkan dengan pengguna di negara tetangga seperti Singapura 66,3% (6 juta jiwa) dan Malaysia 20,4% (34.3 juta jiwa) [3], [4], [5], [6], [7], [8].

Oleh karena kecilnya rasio pengguna *platform LinkedIn* di Indonesia, maka ulasan pengguna *LinkedIn* sangatlah penting untuk diketahui sebab hal ini memberikan prespektif tertentu mengenai pengalaman nyata yang bersifat positif maupun negatif dari berbagai kalangan sesama pengguna *platform linkedin* tersebut.

Dalam penelitian ini, kami telah menjalankan survei terhadap 103 responden, yang dibagi menjadi dua kategori, yaitu mahasiswa dan pekerja. Survei ini untuk mengungkap apakah terdapat korelasi antara rasio pengguna *LinkedIn* di Indonesia dan survey pengguna aplikasi *LinkedIn*.



Gambar 1. 2 Persentase responden menggunakan dan memiliki LinkedIn

Hasil survei menggambarkan bahwa terdapat suatu korelasi, yang menunjukkan bahwa responden dalam pengguna *platform LinkedIn* cenderung jarang menggunakan atau bahkan tidak menggunakan *platform* tersebut. Meskipun demikian, sebanyak 70,9% dari total responden menyatakan bahwa mereka memiliki akun di *LinkedIn*. Sejalan dengan penelitian ini, survei lapangan apakah dapat menggambarkan kesesuaian prespektif pengalaman pengguna *platform LinkedIn* dengan ulasan pengguna *LinkedIn* pada *Google Playstore* di Indonesia.

Untuk melihat ulasan kita dapat mencarinya di *google playstore*. *Google play store* adalah toko aplikasi resmi untuk sistem operasi *Android*, yang memungkinkan pengguna untuk menelusuri dan mengunduh aplikasi yang

dikembangkan dengan *Android software development kit (SDK)* dan diterbitkan melalui *Google* [9].

Untuk mempelajari suatu ulasan (analisis sentimen) secara komputerisasi dibutuhkan keilmuan khusus yaitu cabang dari *Artificial Intelligence (AI)* yang dinamakan *Natural Language Processing (NLP)*. Dengan NLP, memungkinkan komputer menampilkan morfologi, bentuk, serta arti dan kalimat dari suatu bahasa.

Dalam prosesnya, NLP membutuhkan suatu algoritma yang dapat memberikan analisis sentimen terhadap suatu ulasan agar terklasifikasi kedalam positif, netral atau negatif. Algoritma Naïve Bayes adalah algoritma yang sangat populer untuk tugas klasifikasi seperti klasifikasi text. Algoritma ini termasuk dalam keluarga algoritma pembelajaran generatif, yang berarti bahwa tujuannya adalah untuk memodelkan distribusi input terhadap kelas atau kategori tertentu. Berbeda dengan pengklasifikasi diskriminatif, seperti regresi logistik, algoritma ini tidak fokus pada pembelajaran fitur mana yang paling signifikan untuk membedakan antar kelas [10].

Selain menggunakan algoritma Naive Bayes pada penelitian ini juga menggunakan algoritma KNN sebagai pembanding yang dimana KNN adalah Algoritma yang bekerja dengan cara mencari objek data yang jaraknya paling dekat dengan data baru [11]. Objek data yang paling dekat ini kemudian digunakan untuk menentukan kelas data baru. Nilai k menentukan jumlah objek data yang akan digunakan untuk menentukan kelas data baru.

Pada penelitian [12], Penerapan Algoritma K-Nearest Neighbor untuk Analisis Sentimen Terhadap Isu Khilafah dan Radikalisme di Indonesia memiliki akurasi yang didapat sebesar 0,92 dan 0,88. Pada [13], Analisis Sentimen Pada Situs Google Review dengan Naïve Bayes dan Support Vector Machine memiliki akurasi yang didapat sebesar 0,80. Pada [14], Rancang bangun model sentiment analysis review produk pada toko online menggunakan Naïve Bayes memiliki akurasi yang didapat sebesar 0,84. Pada [15], Analisis Sentimen Terhadap Kualitas Pelayanan Aplikasi Gojek Menggunakan Metode Naive Bayes Classifier memiliki akurasi yang didapat sebesar 0,68. Pada [16], Analisis Sentimen Ulasan Pengguna Untuk Peningkatan Layanan Menggunakan Algoritma Naïve Bayes memiliki akurasi yang didapat sebesar 0,83. Pada [17], memiliki akurasi yang didapat sebesar 0,88.

Oleh karena itu, pada penelitian ini penulis melakukan analisis sentimen *platform linkedin* yang telah *discrapping* pada *google play store* dengan rentang waktu September 2018 hingga September 2023 dan mengangkat tema dengan Judul **“Analisis Sentiment Ulasan Pengguna LinkedIn pada Google Play Store dengan Metode Naïve Bayes dan KNN”**.

1.2 Identifikasi Masalah

Adapun identifikasi masalah dari latar belakang diatas yaitu :

1. Rasio pengguna *Linkedin* di Indonesia masih sedikit dibandingkan dengan Malaysia dan Singapura
2. Klasifikasi sentimen dengan menggunakan Naïve Bayes dan KNN

1.3 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan diatas, terdapat beberapa masalah yang dapat terjadi, maka rumusan masalah yang akan dibahas sebagai berikut :

1. Bagaimana gambaran ulasan pengguna *Linkedin* yang ada di Indonesia ?
2. Bagaimana tingkat akurasi Algoritma Naïve bayes dan KNN pada sentimen analisis Ulasan *Linkedin* ?

1.4 Batasan Masalah

Batasan masalah dalam penelitian ini sebagai berikut :

1. Data yang diambil berasal dari ulasan *Linkedin* pada *Google Playstore*
2. Jumlah ulasan data yang di *Scrapping* sebanyak 2.000 ulasan data
3. Bahasa Asing dan Bahasa daerah diabaikan
4. Klasifikasi sentiment ulasan *LinkedIn* dibagi menjadi 2, yaitu sentiment positif dan negative menggunakan metode Naïve Bayes dan KNN
5. Proses Text Mining melibatkan *casefolding*, *normalize*, *stopwords*, *tokenize*, dan *stemming*
6. Ekstrasi fitur menggunakan TF – IDF
7. Evaluasi performa model menggunakan *confusion matrix*

1.5 Tujuan Penelitian

Penelitian ini bertujuan untuk :

1. Mengklasifikasikan apakah algoritma Naïve Bayes dan KNN dapat digunakan untuk proses klasifikasi ulasan pengguna aplikasi *Linkedin* pada aplikasi *Google Playstore*

2. Mengetahui faktor – faktor sentimen atau opini pengguna *platform LinkedIn*
3. Seberapa besar keakuratan algoritma Naïve bayes dan KNN pada klasifikasi ulasan pengguna LinkedIn

1.6 **Manfaat Penelitian**

Penelitian ini Bermanfaat untuk :

1. Mahasiswa dapat menerapkan dan mengetahui Algoritma Naïve bayes dan KNN pada Sentiment Analisis
2. Mahasiswa dapat mengetahui ke akuratan algoritma Naïve Bayes dan KNN pada Sentiment Analisis

1.7 **Sistem Penulisan**

Untuk memberikan gambaran yang jelas mengenai penelitian ini, dibuatlah sistematika penulisan yang terdiri dari beberapa bab sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang masalah, identifikasi masalah, rumusan masalah, batasan masalah, tujuan, dan manfaat penelitian

BAB II LANDASAN TEORI

Pada bab ini, akan dibahas berbagai teori pendukung dan landasan teoritis yang relevan dengan topik yang diselidiki.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan tentang penelitian serta metode yang digunakan dalam penyelidikan materi pelajaran.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini, penulis akan mendiskusikan hasil dan temuan dari penelitian yang telah dilakukan.

BAB V PENUTUP

Bab ini berisi tentang kesimpulan dan rekomendasi dari penelitian yang diharapkan dapat memberikan manfaat bagi pembaca.

