

BAB I

PENDAHULUAN

1.1 Latar Belakang

Social network merupakan sebuah platform berbasis *website* ataupun aplikasi *mobile*, yang memungkinkan penggunanya membuat profil yang dapat diakses oleh publik baik secara terbuka maupun secara privat. *Social network* memungkinkan untuk berkomunikasi kepada sesama pengguna serta berbagi konten digital dalam bentuk teks, gambar, maupun video (Boyd & Ellison, 2007).

Twitter merupakan situs *social network* yang penggunanya dapat mengirim, berbagi dan membalas pesan singkat atau yang biasa disebut tweet. Layanan *microblogging* ini digemari penggunanya karena akses informasi *real-time* dan komunikasi yang mudah dengan teman, akun yang dikelola *brand*, *business leaders*, politikus, serta selebriti. Atas fitur unggulannya itulah Twitter memiliki pengguna loyalnya. Selain itu, Twitter memiliki keunikan pada tweet-nya, yaitu permanen dan dapat dicari melalui *search engine* secara bebas oleh publik. Adapun opsi untuk melindungi tweet sehingga hanya pengikut pengguna yang dapat membaca *tweet* (Hetler, 2022).

Sejak beberapa tahun terakhir, pengguna Twitter di Indonesia terus bertumbuh pesat, bahkan Indonesia menjadi negara kelima dengan pengguna Twitter terbanyak di dunia. Berdasarkan StatCounter, terdapat 4.45% pengguna Twitter dari total market share media sosial di Indonesia yang mencapai angka 170 juta.

Menurut Febriani dan Dewi, pengguna media sosial di Indonesia akan mencari informasi sebanyak-banyaknya tentang barang atau jasa untuk

menghilangkan kebingungan sebelum membeli barang dan mengantisipasi penyesalan yang akan dialami apabila keputusan pembelian tidak tepat (Febriani & Dewi, 2019). Salah satu unsur yang terdapat pada media sosial yang membuat banyak *brand* memutuskan untuk menjalankan kampanye *brand*-nya melalui media sosial adalah *electronic word of mouth* (eWOM). eWOM adalah proses pertukaran informasi yang dinamis antara calon pelanggan yang potensial, pelanggan yang sedang bertransaksi, dan pelanggan yang pernah bertransaksi tentang produk, jasa, *brand*, atau perusahaan, yang dapat diakses khalayak luas melalui internet (Ismagilova dkk. 2017). Keberadaan eWOM mempengaruhi pengguna media sosial dalam menentukan pilihan sampai pada tingkatan tertentu, salah satu variabel yang akan meningkatkan tingkat persuasi dari kutipan eWOM di media sosial adalah narasumbernya (Luo dkk. 2018). Narasumber eWOM lebih akan mungkin dipercaya apabila narasumber tersebut dikenal oleh banyak orang atau memiliki reputasi yang baik. Akan tetapi, dalam penelitiannya, Luo menambahkan, sesuai sifat dari internet, pengguna tidak memiliki cukup sumber daya waktu untuk dapat menilai keprofesionalan narasumber dan keaslian informasinya, sehingga berdampak pada kesalahan pengguna dalam pengambilan keputusan.

Strategi pemasaran dengan pendekatan eWOM merupakan angin segar dalam bidang pemasaran untuk saat ini. Sebelumnya, praktisi pemasaran berkompetisi dengan ide-ide organik dan otentik mereka. Namun dengan pendekatan eWOM, tim pemasaran mulai memiliki banyak opsi dalam strategi pemasarannya. Salah satu strateginya adalah dengan menggunakan konten fabrikasi yang disebar dengan jasa *buzzer*. Konten fabrikasi menurut Ireton dan Posseti (2017) dikategorikan sebagai disinformasi. Disinformasi merupakan informasi

yang salah. Kebenaran dari konten fabrikasi tentu saja diketahui oleh penyebar, penyebar tahu bahwa informasi yang disebarkan merupakan informasi yang tidak benar. Penyebaran konten fabrikasi dilakukan oleh aktor yang secara sengaja membuat dan menyebarkannya ke publik, aktor ini dikenal dengan istilah *buzzer*. Juzar dan Akbar memaparkan bahwa *buzzer* adalah pengguna media sosial yang memiliki pengaruh dan dapat mempengaruhi pengguna lain. Maulana dan Kuswayati menambahkan bahwa konten dari *buzzer* dapat difabrikasi dengan kesepakatan nominal yang telah ditentukan. Dengan konten yang difabrikasi dan diedarkan, *buzzer* bayaran dapat menjadi sosok yang dapat diandalkan untuk pertumbuhan tingkat engagement yang murah serta kecil kemungkinan berdampak negatif. Pihak yang menggunakan jasa *buzzer* dapat diuntungkan sesuai dengan sifat eWOM yaitu mempengaruhi.

Dengan maraknya *campaign* terfabrikasi di Twitter yang bertujuan untuk meningkatkan sentimen positif dari *brand*, banyak pengguna yang awalnya mempercayakan dasar penentuan keputusannya berdasarkan eWOM, justru menyesali pilihannya dan merasa ditipu dikarenakan mereka terlanjur mempercayai eWOM positif yang difabrikasi oleh *buzzer* tersebut.

Korban kerugian eWOM terfabrikasi yang cukup mendapat perhatian adalah yang berkaitan dengan investasi terhadap *e-commerce* buatan Indonesia. Selain itu, tentu saja masih banyak intensi lain dari kampanye terfabrikasi lainnya dengan cara menyebarkan *tweets* positif di Twitter. Dikarenakan sifat *tweet* yang unik serta rumitnya dalam menentukan nilai dan keaslian. Untuk itu, perlu adanya bantuan yang dapat mengidentifikasi akun penyebar eWOM (*buzzer*) dengan akurat dan efisien, salah satunya dengan pendekatan data mining. Data mining merupakan

proses menemukan struktur yang menarik dalam data. Struktur yang dapat mengambil banyak bentuk, termasuk seperangkat aturan, sebuah grafik atau jaringan, sebuah tree, satu atau beberapa persamaan, dan banyak lagi. Struktur dapat menjadi bagian dari dasbor visual yang kompleks atau sederhana (Roiger, 2017).

Pada penelitian identifikasi akun *buzzer* yang dilakukan oleh Maulana dan Kusmayati, frekuensi *tweet* harian, usia akun sejak pembuatan, banyaknya URL yang disematkan dalam *tweet*, banyaknya *retweet* yang dilakukan dalam *tweet*, banyaknya follower, serta banyaknya following menjadi atribut yang digunakan pada pengklasifikasian. Pada penelitian ini, algoritma Naive Bayes akan digunakan sebagai basis model algoritma untuk membantu pengklasifikasian terhadap akun Twitter, apakah dikategorikan sebagai *buzzer* atau akun asli yang dikelola manusia.

Naive Bayes sendiri merupakan algoritma dalam *machine learning* yang dapat memperkirakan probabilitas suatu kejadian terjadi berdasarkan kejadian lain yang pernah terjadi. Alasan Naive Bayes digunakan sebagai metode penelitian yaitu untuk memprediksi probabilitas. Bekerja dengan menggunakan konsep mixture models. Mixture model dapat menentukan probabilitas komponen yang ada pada teorema Bayes untuk bertindak sebagai pengklasifikasi probabilitas (Rahat dkk, 2019).

Varian yang digunakan dalam penelitian ini merupakan varian Gaussian Naive Bayes. Gaussian Naive Bayes merupakan salah satu varian dari Naive Bayes yang menerapkan Gaussian *normal distribution*, varian ini mendukung data *continuous* untuk klasifikasi (Ismail dkk, 2020). Gaussian Naive Bayes dengan karakteristiknya memiliki pendekatan yang sedikit berbeda dan dapat digunakan dengan efisien pada klasifikasi biner (Sushma dan Keerthan, 2021).

Pada penelitian ini, akan diungkap *buzzer* dengan cara mempelajari *pattern* dari akun-akun yang ikut menyebarkan kampanye pada *Initial Public Offering* (IPO) perusahaan *e-commerce* di Indonesia yang sempat *trending* pada Twitter. Dimulai dengan pengumpulan data dari fenomena yang terjadi, pembersihan data, pembobotan sentimen, pelabelan manual sesuai dengan karakteristik *buzzer* yang telah didefinisikan, dan diakhiri dengan pengujian klasifikasi data terhadap model.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang sudah dijabarkan sebelumnya maka dapat dirumuskan beberapa permasalahan seperti:

1. Apa pendekatan yang perlu dilakukan untuk mengetahui suatu akun penyebar ulasan terhadap topik yang sedang *trending* atau ramai dibicarakan merupakan *buzzer* atau organik?
2. Apa saja atribut yang digunakan dalam klasifikasi yang dapat membantu mengidentifikasi keaslian akun?
3. Bagaimana hasil dari penerapan algoritma Naive Bayes pada pengklasifikasi *buzzer*?

1.3 Batasan Masalah

Dalam penelitian ini terdapat beberapa batasan masalah, seperti :

1. Penelitian ini tidak mencakup *buzzer* yang memiliki skor *engagement* yang tinggi seperti selebriti atau *influencer*.
2. Rentang waktu dari data yang digunakan dalam penelitian yaitu mulai dari 10 Juni 2021 sampai dengan 9 Desember 2022.
3. Data yang digunakan dalam *training* akun *buzzer* merupakan data yang didapatkan melalui Twitter developer menggunakan *library* pendukung

terkait yang berupa *tweet* pada topik IPO *e-commerce*, *tweet* pada akun-akun yang terlibat pada topik tersebut (hingga 10 *tweet*), umur akun, jumlah *following*, dan jumlah *followers*.

4. Bahasa pemrograman yang digunakan adalah *python* dan menggunakan lingkungan pengembangan Google Colab.

1.4 Tujuan

Berdasarkan rumusan masalah yang telah dipaparkan sebelumnya maka didapatkan tujuan penelitian sebagai berikut:

1. Mengidentifikasi keaslian akun yang mengedarkan ulasan-ulasan bersentimen positif.
2. Membantu memastikan keaslian eWOM dari *tweet* sehingga dapat digunakan sebagai dasar menentukan keputusan karena keaslian penyebar eWOM-nya dapat dibuktikan.
3. Mengungkap bahwa praktik *buzzer* dapat diketahui polanya sehingga dapat menjadi *insight* bagi organisasi bisnis untuk berhenti menggunakan *buzzer*.

1.5 Sistematika Penulisan

Sistematika penulisan skripsi ini bertujuan untuk membantu penyusunan laporan skripsi agar sesuai dengan acuan yang telah ditentukan dan mencapai tujuan penulisan yang telah ditetapkan. Beberapa tahapan yang harus dilalui dalam proses penyusunan skripsi ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini menjelaskan gambaran umum tentang penelitian yang akan dilakukan, termasuk latar belakang penelitian, rumusan masalah

yang akan dipecahkan, batasan masalah yang ditetapkan, tujuan penelitian, dan sistematika penulisan yang akan digunakan.

BAB II TINJAUAN PUSTAKA

Bab ini berisi tentang daftar penelitian yang sudah pernah dilakukan sebelumnya yang bertujuan untuk menjadi dasar penelitian. Penelitian yang dibahas adalah penelitian yang memiliki keterkaitan dengan penelitian ini seperti analisis sentimen, *buzzer*, *text mining*, klasifikasi, analisis sentimen, Twitter, Vader, Social Network Scrape, Pandas, Naive Bayes, dan Sklearn.

BAB III METODOLOGI PENELITIAN

Bab ini berisi tentang metode penelitian yang akan dibuat yaitu meliputi studi literatur, analisis kebutuhan, dan perancangan model dan visualisasi.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan proses pembangunan, pelatihan, dan pengujian klasifikasi secara lebih terperinci, serta menampilkan hasil yang diperoleh dari klasifikasi tersebut.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari seluruh hasil penelitian yang telah dilakukan dan saran terhadap penelitian selanjutnya.

DAFTAR PUSTAKA

Berisi tentang literatur yang digunakan sebagai pedoman yang membantu pengerjaan penelitian.

LAMPIRAN

Pada bagian ini berisi tentang data atau pelengkap yang menunjang dalam pembuatan skripsi.