

## **BAB I**

### **PENDAHULUAN**

Dalam bab ini akan dibahas latar belakang dilaksanakannya penelitian, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

#### **1.1 Latar Belakang**

Pada saat ini, media sosial sudah digunakan oleh sebagian besar masyarakat Indonesia. Berdasarkan data dari We Are Social (2023) Indonesia memiliki pengguna aktif media sosial mencapai 167 juta pengguna per Januari 2023. Artinya 60,4% dari populasi dalam negeri Indonesia menggunakan media sosial. Banyaknya pengguna tersebut membuat media sosial memiliki berbagai peran bagi penggunanya. Salah satu peran yang paling penting adalah sebagai sarana untuk memperoleh informasi.

Twitter merupakan salah satu media sosial yang banyak digunakan di Indonesia dengan jumlah pengguna mencapai 25,25 juta pada tahun 2023 (We Are Social, 2023). Twitter adalah media sosial berbentuk *microblog* yang membuat penggunanya dapat mengirimkan dan membaca informasi secara *real time* sehingga dapat menyebarkan segala jenis informasi dengan masif dalam waktu yang singkat, tidak terkecuali informasi palsu atau *hoax*.

*Hoax* adalah pemberitaan palsu yang bertujuan untuk membohongi orang lain. Penyebaran *hoax* merupakan tindakan melanggar hukum yang diatur dalam Pasal 45A ayat (1) Undang-Undang Nomor 19 Tahun 2016 tentang Perubahan atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik (UU ITE 2016) (Indonesia) dengan ancaman pidana penjara paling lama 6 (enam) tahun dan/atau denda paling banyak Rp 1.000.000.000,00 (satu miliar rupiah).

Penyebaran *hoax* di media sosial dapat memberikan masalah yang serius. Salah satu kasus yang terjadi akibat dari penyebaran *hoax* di sosial media adalah kerusuhan yang terjadi di Manokwari, Papua Barat pada hari Senin tanggal 19 Agustus 2019. Pada kerusuhan tersebut, terjadi pembakaran sejumlah bangunan, yaitu rumah warga hingga Gedung DPRD

Papua Barat. Kerusuhan tersebut disebabkan oleh penyebaran *hoax* foto mahasiswa yang tewas disebabkan kejadian di Jawa Timur (Movanita, 2019). Informasi palsu yang disebarakan tersebut membuat warga Manokwari Papua Barat menjadi emosi dan melakukan aksi kerusuhan di wilayahnya (Movanita, 2019).

Salah satu pelaku yang paling sering menyebarkan *hoax* di media sosial adalah *buzzer* (Juditha, 2019). *Buzzer* awalnya merupakan sebuah teknik pemasaran barang atau jasa untuk menghasilkan bisnis dengan pergerakan informasi dari mulut ke mulut atau disebut juga dengan *buzz marketing* (Mustika, 2019). *Buzz marketing* merupakan upaya penguatan pemasaran oleh pihak ketiga, baik secara pasif maupun aktif yang ditujukan untuk saling memengaruhi (Thomas Jr, 2004). *Buzzer* dianggap sebagai salah satu aktor yang paling penting dalam penggalangan opini di media sosial dalam menjalankan pemasaran (Juditha, 2019).

Saat ini, istilah *buzzer* di Indonesia mulai mengalami pergeseran menjadi konsep yang memiliki stereotip negatif (Sugiono, 2020). Hal tersebut disebabkan oleh kegiatan tidak bertanggung jawab yang dilakukan oleh *buzzer* politik profesional di media sosial (Felicia, 2018). Keterlibatan *buzzer* dalam peristiwa politik di Indonesia telah membuat kontribusi negatif terhadap citra dan pemaknaan masyarakat terhadap *buzzer* yang dianggap sebagai pihak yang dibayar untuk memproduksi konten negatif di media sosial (Mustika, 2019). *Buzzer* merupakan aktor yang banyak menyebabkan terjadinya fenomena *hoax*, ujaran kebencian, fitnah dan kampanye negatif lainnya terus bertumbuh akibat penyebaran pesan-pesan yang dilakukan oleh para *buzzer* (Juditha, 2019). *Buzzer* dapat dengan mudah menggeser perdebatan ke hal-hal yang tidak substansial, seperti menyeragamkan narasi serta memutarbalikkan fakta sehingga perhatian publik akan teralihkan dari sebuah isu penting (Mustika, 2019).

Berdasarkan potensi dari dampak negatif yang ditimbulkan oleh *buzzer*, deteksi akun-akun yang berperan sebagai *buzzer* di media sosial merupakan hal yang sangat penting (Afra dkk, 2024). Selain itu, deteksi akun *buzzer* diperlukan karena data di media sosial yang berasal dari akun *buzzer* dianggap sebagai *noise* dalam penelitian-penelitian lain, seperti analisis sentimen sehingga data tersebut perlu dihilangkan (Suciati dkk, 2019).

Deteksi *buzzer* dapat dilakukan secara otomatis dengan memanfaatkan *machine learning*. *Machine learning* tersebut digunakan untuk membangun model klasifikasi yang dapat mendeteksi akun-akun yang termasuk dalam kategori akun *buzzer* pada media sosial Twitter. Model deteksi tersebut dibangun dengan mengimplementasikan dua algoritma, yaitu *XGBoost* dan *Naive Bayes*. Algoritma *XGBoost* memiliki keunggulan dalam dalam hal kinerja, efisiensi waktu, dan memori (Wang dkk, 2020). Sementara itu, Algoritma *Naive Bayes* memiliki beberapa keunggulan, yaitu cepat, mudah diimplementasikan dengan struktur sederhana, dan efektif (Taheri dan Mammadov, 2013). Penelitian ini bertujuan untuk mengetahui kinerja dari kedua algoritma tersebut dalam mendeteksi akun *buzzer* dan *non-buzzer* pada media sosial Twitter.

Beberapa penelitian lain telah dilakukan untuk mendeteksi akun *buzzer* di media sosial. Suciati dkk (2019) mendeteksi *buzzer* pada media sosial Twitter menggunakan empat model *machine learning*, yaitu *Ada Boost (AB)*, *Gradient Boosting (GB)*, *Extreme Gradient Boosting (XGB)*, and *Histogrambased Gradient Boosting (HGB)*. Fitur-fitur yang digunakan dalam penelitian ini merupakan fitur yang berbasis pada data masing-masing akun, seperti lokasi akun tersebut, jumlah pengikut akun dan jumlah *tweet* akun.

Penelitian lain dilakukan oleh Ibrahim dkk (2015) untuk mendeteksi akun-akun *buzzer* di media sosial Twitter pada masa pemilihan presiden. Penelitian tersebut menggunakan pendekatan *machine learning* dengan fitur-fitur yang berbasis pada data masing-masing akun, seperti usia akun, jumlah *retweet*, dan jumlah pengikut.

Penelitian yang telah dilakukan sebelumnya menggunakan pendekatan berbasis data pada masing-masing akun dan tidak menjelaskan hubungan diantara akun-akun tersebut. Penelitian ini menggunakan pendekatan berbeda, yaitu menggunakan *social network analysis* untuk mendeteksi akun yang berperan sebagai *buzzer* dengan mengidentifikasi pola-pola pertemanan dan *retweet buzzer* dalam jaringan sosial.

*Social network analysis* merupakan sebuah studi yang mempelajari tentang hubungan antar manusia dengan memanfaatkan teori graf (Tsvetovat & Kouznetsov, 2011:1). *Social network analysis* melakukan pemanfaatan teori graf dengan memandang hubungan sosial sebagai *node* yang merupakan aktor dalam suatu jaringan dan *edge* yang merupakan

hubungan antar aktor dalam jaringan tersebut. *Social network analysis* menjelaskan pola hubungan yang terdapat di antara aktor serta menggambarkan bagaimana hubungan dalam jaringan tersebut (Oktora & Alamsyah, 2014).

*Social network analysis* dapat digunakan untuk mengetahui struktur dan jaringan *buzzer* di media sosial (Camil dkk, 2017). Proses *social network analysis* menghasilkan metrik-metrik berdasarkan posisi akun dalam jaringan pertemanan dan *retweet*. Metrik-metrik tersebut akan digunakan sebagai fitur dalam pembangunan model klasifikasi *machine learning* untuk mendeteksi akun-akun *buzzer*. Penelitian ini bertujuan untuk mengetahui kinerja dari metrik-metrik *social network analysis* yang digunakan sebagai fitur dalam model deteksi yang dibangun.

Data yang digunakan pada penelitian ini berasal dari sebuah topik yang terdapat pada media sosial Twitter. Topik tersebut merupakan topik buatan atau topik yang secara sengaja dibuat oleh akun-akun *buzzer* sehingga mayoritas akun yang terlibat pada topik ini berasal dari kalangan *buzzer*. Dominasi akun *buzzer* tersebut membuat deteksi yang dilakukan harus memiliki kemampuan yang baik dalam menghindari kesalahan mendeteksi akun dari kalangan pengguna biasa yang dikategorikan sebagai akun *buzzer*. Hal tersebut dapat bermanfaat ketika pihak Twitter memutuskan untuk memblokir secara massal akun-akun *buzzer* yang berperan dalam pembuatan topik buatan, maka tidak ada akun pengguna biasa yang ikut terblokir yang dapat merugikan pengguna tersebut.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka dirumuskan masalah sebagai berikut:

1. Bagaimana kinerja algoritma *machine learning* *XGBoost* dan *Naive Bayes* dalam mendeteksi akun *buzzer* dan *non-buzzer* pada media sosial Twitter?
2. Bagaimana kinerja metrik-metrik *social network analysis* yang digunakan fitur dalam model *machine learning*?

### 1.3 Batasan Masalah

Untuk memfokuskan penelitian, ditetapkan beberapa batasan masalah, yaitu sebagai berikut:

1. Data yang digunakan dalam penelitian ini merupakan data akun yang menulis *tweet* dengan kata kunci “KALAHkan TukangFITNAH” yang dikumpulkan pada tanggal 12 Februari 2024 sampai 13 Februari 2024 serta akun yang memberikan balasan terhadap *tweet* tersebut.
2. Penelitian ini berfokus pada pembahasan komparasi kinerja model deteksi yang dibangun menggunakan algoritma *XGBoost* dan *Naive Bayes*.
3. Penelitian ini berfokus pada pembahasan komparasi kinerja fitur yang digunakan dalam model deteksi.

### 1.4 Tujuan Penelitian

Berdasarkan rumusan masalah yang ada, maka tujuan penelitian ini yaitu sebagai berikut:

1. Mengetahui kinerja algoritma *machine learning XGBoost* dan *Naive Bayes* dalam model deteksi akun *buzzer* dan *non-buzzer* yang telah dibangun.
2. Mengetahui kinerja metrik-metrik *social network analysis* yang digunakan fitur dalam model *machine learning*.

### 1.5 Manfaat Penelitian

Dengan dilakukannya penelitian ini diharapkan dapat memberikan manfaat sebagai berikut:

1. Memberikan data dan komparasi dari berbagai algoritma mengenai deteksi akun *buzzer* dan *non-buzzer* pada media sosial Twitter.
2. Memberikan data dan komparasi dari penggunaan metrik-metrik *social network analysis* sebagai fitur dalam model deteksi akun *buzzer* dan *non-buzzer* pada media sosial Twitter.

3. Dapat menjadi rujukan bagi penelitian selanjutnya yang memiliki keterkaitan dengan penelitian ini baik secara langsung maupun tidak langsung.

## **1.6 Sistematika Penulisan**

Adapun sistematika penulisan karya ilmiah ini adalah sebagai berikut:

### **BAB I PENDAHULUAN**

Bab ini berisi latar belakang, rumusan masalah, batasan masalah, tujuan penelitian yang akan dilakukan, dan sistematika penulisan.

### **BAB II TINJAUAN PUSTAKA**

Bab ini berisi penjelasan tentang teori-teori dan konsep algoritma yang digunakan dalam penelitian.

### **BAB III METODOLOGI PENELITIAN**

Bab ini berisi penjelasan langkah-langkah yang akan dilakukan dalam penelitian.

### **BAB IV HASIL PENELITIAN DAN PEMBAHASAN**

Bab ini berisi uraian tentang hasil penelitian dan pembahasan terhadap hasil penelitian yang dilakukan.

### **BAB V KESIMPULAN DAN SARAN**

Bab ini berisi kesimpulan dari keseluruhan penelitian yang telah dilakukan, serta saran dari penulis untuk kegiatan penelitian selanjutnya terkait dengan topik yang sedang dibahas.