

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Menurut Kelana (2011), Alay merupakan singkatan dari “anak layangan” atau “anak lebay”. Alay menggambarkan gaya hidup yang norak, kampungan dan berlebihan untuk menarik perhatian orang disekitarnya. Gaya hidup tersebut di alami oleh remaja Indonesia yang ingin selalu terlihat eksis atau narsis mereka dalam segala hal misalnya berpakaian, perilaku, dan berbahasa baik lisan maupun tulisan. Mereka berpakaian dengan motif atau warna yang terlihat sangat jelas misalnya menggunakan baju berwarna merah dan celana yang berwarna biru. Mereka juga menggunakan celana jeans yang ketat dengan tatanan rambut seperti Kangen Band atau Syahrini (Kelana, 2011).

Bahasa Alay merupakan variasi bahasa Indonesia yang trend digunakan oleh kalangan para remaja untuk berkomunikasi antara satu dengan yang lain. Bahasa Alay memiliki struktur yang berbeda dengan struktur bahasa baku (Hermaji, 2014). Terdapat 5 ciri-ciri bahasa Alay yaitu, kombinasi huruf besar dan huruf kecil, kombinasi huruf dan angka, menggunakan huruf yang sama dan berulang, menggunakan istilah-istilah alay atau kombinasi semuanya (Pardosi, 2014). Berdasarkan ciri-ciri tersebut dapat dilihat bahwa penulisan bahasa Alay itu beragam, unik dan tidak beraturan. Belum dapat dipastikan kapan bahasa Alay ini mulai menjamur di kalangan para remaja, namun tanda-tanda bahasa Alay sudah ada sejak lama dan tanpa disadari mungkin kita juga pernah menggunakannya pada saat remaja atau pun di saat sekarang (Kelana, 2011).

Pengguna bahasa Alay adalah remaja, namun setiap remaja belum tentu Alay (Kelana, 2011). (Hermawan, 2016) menyatakan Masa remaja antara usia 17 -24 tahun untuk wanita dan usia 17 – 25 tahun untuk pria. Pada media sosial sulit untuk mengetahui rentang usia pada akun Twitter seseorang, karena tidak semua pengguna Twitter menampilkan tanggal lahir pada akunnya. Dengan demikian sulit untuk dapat membedakan pengguna Twitter tersebut remaja atau bukan, dan hal

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

tersebut dapat diketahui dari gaya bahasa yang digunakan yaitu bahasa Alay (Kelana, 2011).

Penggunaan bahasa Alay berkembang karena adanya SMS (*Short Message Service*) atau pesan singkat dari layanan operator yang mengenakan tarif per karakter atas per SMS dengan tujuan untuk menghemat biaya. Perkembangan teknologi komunikasi membuat pesan singkat semakin banyak ditemui di jejaring sosial seperti *Facebook*, *Twitter*, dan *Instagram*.

*Twitter* dipilih karena, merupakan salah satu media sosial yang banyak diminati masyarakat, tahun 2017 Indonesia menempati urutan ke lima setelah USA, Brazil, Jepang, dan Inggris sebagai pengguna *Twitter* terbanyak di dunia dengan total 328 juta pengguna. Jumlah pengguna aktif *Twitter* mencapai 24,34 juta. Beberapa istilah yang terdapat pada *Twitter* yaitu *direct message*, *favorite*, *follow*, *unfollow*, *follower*, *following*, *hashtag*, *retweet*, *timeline* dan *tweet*. Alasan lain kenapa *twitter*, karena *tweet* pada *Twitter* dibatasi maksimal hanya 140 karakter, yang memungkinkan para pengguna *Twitter* melakukan penyingkatan secara Alay. *Tweet* dapat di lihat secara publik, namun pengirim dapat membatasi pengiriman pesan (*direct message*) ke daftar *follower* mereka saja. Pengguna juga dapat melihat *tweet* dari pengguna lain.

Menurut Jaelani dan Melisa (2014), variasi bahasa yang sering digunakan untuk menuliskan *tweet* adalah bahasa Alay. Hal ini juga terlihat berdasarkan *dataset* yang dikumpulkan oleh Hidayatullah (2015) terkait karakteristik *tweet* oleh pengguna yang berasal dari Indonesia, dimana sebanyak 85,40% mengandung kata yang ditulis secara Alay, lebih banyak dibandingkan penggunaan singkatan dan akronim, kata dalam bahasa asing dan bahasa daerah. Penelitian-penelitian tersebut memperlihatkan banyaknya pengguna *Twitter* di Indonesia yang menggunakan bahasa Alay.

Menurut (Maulina dan Sumarni, 2015), perilaku anak Alay di media sosial *Twitter* yaitu menulis *tweet* dengan ciri khas penulisan anak Alay, terlalu sering meng-*update* status di jejaring sosial, dan sering meng-*upload* foto dengan gaya dua jari, atau dengan gaya dari berbagai sisi misalnya gaya yang ke samping atau ke atas. Foto yang di-*upload* juga telah diubah agar terlihat lebih cantik. Nama akun

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

pengguna Twitter biasanya menggunakan nama samaran, bukan nama sendiri. Ini menunjukkan bahwa mereka ingin terlihat lebih baik lebih hebat dan lebih keren di bandingkan dengan yang lain.

Untuk memudahkan identifikasi akun Alay di jejaring sosial, dan untuk lebih memahami karakteristik anak Alay ketika menggunakan jejaring sosial, dalam penelitian ini akan dilakukan pengklasifikasian akun di Twitter ke dalam dua kelas, yaitu Alay dan Tidak Alay. Uddin, Imran, dan Sajjad (2014), telah melakukan pengklasifikasian akun Twitter ke dalam kelas *real actor* (manusia) dan *digital actor* (program komputer, *bot*). Mereka menggunakan 17 fitur berdasarkan informasi yang diperoleh dari profil Twitter dan perilaku *tweet*-nya. Dalam penelitian tugas akhir ini, fitur-fitur tersebut akan digunakan dan diperluas sesuai dengan permasalahan Alay di jejaring sosial.

Untuk mengklasifikasikan akun Alay di Twitter, akan dipergunakan metode *Naïve Bayes Classifier*. NBC merupakan salah satu metode dalam *text mining* yang digunakan untuk mengklasifikasikan sebuah data dengan menggunakan perhitungan probabilistik dan statistik (Thomas Bayes). *Naïve Bayes Classifier* dipilih karena sederhana, proses pengolahan datanya mudah dan memberikan tingkat akurasi yang baik.

Nagpal (2014) membandingkan tiga metode *decision tree*, SVM dan *Naïve Bayes Classifier* untuk mengklasifikasikan pengguna Twitter, dan *Naïve Bayes Classifier* merupakan metode yang paling tinggi tingkat akurasinya yaitu 80,9%, dibandingkan dengan SVM dan *decision tree*. Liantoni (2015), mengklasifikasi daun herbal dengan menggunakan metode *Naïve Bayes* dan *K-Nearest Neighbor* dengan tingkat akurasi *Naïve Bayes Classifier* yaitu 75% lebih tinggi dibandingkan dengan tingkat akurasi KNN yaitu 70,83%. Penelitian (Matsumoto dkk (2016) berjudul *Slang Analysis Based on Variant Information Extraction*

*Focusing on the Time Series Topics*, menggunakan metode *Naïve Bayes Classifier* dengan nilai akurasi yang didapat 82,145%. Ugheoke, (2014) mendeteksi jenis kelamin pengirim *tweet* dengan menggunakan metode *Naïve Bayes Classifier* memperoleh tingkat akurasi 86,8% dan 95,4%.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## 1.2 Rumusan Masalah

Pada tugas akhir ini, dapat dirumuskan permasalahan yaitu bagaimana mengklasifikasikan pengguna Twitter ke dalam dua kelas yaitu Alay atau Tidak Alay dengan menggunakan metode *Naïve Bayes Classifier*.

## 1.3 Batasan Masalah

Pada tugas akhir ini, terdapat sejumlah batasan masalah yaitu :

1. Jumlah *tweet* diambil dari 50 akun yang berasal dari Indonesia, di mana setiap satu akun Twitter minimal telah membuat *tweet* sebanyak 100 *tweet*.
2. Pengklasifikasian didasarkan pada informasi *teks* yang diperoleh profil Twitter dan perilaku *tweet*-nya.
3. Fitur-fitur yang digunakan untuk pengklasifikasian berdasarkan 17 fitur pada Muhammad Moeen Uddin, dkk (2014) dengan penyesuaian dan penambahan sesuai permasalahan Alay di jejaring sosial, sehingga diperoleh sebanyak 18 fitur, yaitu : *follower, following, retweet, name, username, description, liked, like, plain status, mention, URL, hashtag, life time, tweet spread, user collective activeness, collective influence*, jumlah foto/video dan *emoticon*.

## 1.4 Tujuan Penelitian

Tujuan dibuatnya tugas akhir ini yaitu sebagai berikut :

1. Mengklasifikasikan pengguna Twitter dengan menggunakan metode *Naïve Bayes Classifier* ke dalam kelas Alay dan Tidak Alay.
2. Merancang bangun aplikasi klasifikasi akun Alay di Twitter.
3. Menguji akurasi metode *Naïve Bayes Classifier* dengan menggunakan *Confusion Matrix* pada aplikasi klasifikasi akun Alay di Twitter.

## 1.5 Sistematika Penulisan

Penulisan tugas akhir ini terdiri atas beberapa bagian utama yaitu :

### BAB I PENDAHULUAN

Bab pertama membahas tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian dan sistematika penulisan tugas akhir.

**Hak Cipta Dilindungi Undang-Undang**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

## **BAB II LANDASAN TEORI**

Bab kedua membahas tentang bahasa Alay, proses *text mining*, dan metode *Naïve Bayes Classifier*.

## **BAB III METODOLOGI PENELITIAN**

Pada bab ketiga membahas langkah-langkah yang dilakukan dalam proses penelitian, yaitu perumusan masalah, studi pustaka, pengumpulan data dan informasi, analisa, pengujian serta kesimpulan dan saran.

## **BAB IV ANALISIS DAN PERANCANGAN**

Bab keempat membahas tentang analisa dan perancangan sistem yang akan dibuat pada penelitian klasifikasi akun Alay di Twitter dengan menggunakan metode *Naïve Bayes Classifier*.

## **BAB V IMPLEMENTASI DAN PENGUJIAN**

Bab kelima membahas tentang implementasi dari hasil analisa dan perancangan sistem, serta melakukan pengujian untuk mengetahui apakah hasil implementasi telah sesuai dengan yang diharapkan atau tidak.

## **BAB VI KESIMPULAN DAN SARAN**

Bab terakhir ini berisikan kesimpulan dari penelitian yang telah dilaksanakan dan saran perbaikan dan penyempurnaan kedepannya.