

BAB I PENDAHULUAN

1.1 Latar Belakang

Saat ini sosial media menjadi media yang sangat populer untuk mendapatkan berbagai informasi bagi orang-orang dari berbagai kalangan[1]. Banyak informasi penting atau peristiwa terkini yang dapat diakses secara mudah dan cepat menggunakan sosial media. Seiring perkembangan sosial media dewasa ini, *Twitter* menjadi media yang banyak digunakan oleh masyarakat di Indonesia tak terkecuali oleh para Tokoh Publik Politik[2].

Tokoh Publik Politik di Indonesia menggunakan *Twitter* sebagai media untuk berinteraksi dengan masyarakat. Ada berbagai informasi, opini atau pendapat, kritik dan saran yang mereka posting setiap harinya. Dengan banyaknya topik yang sering dibahas dalam postingan *tweet* milik Tokoh Publik Politik di Indonesia, diperlukan waktu yang lebih untuk menentukan apa saja topik yang paling sering dibahas dan menjadi *trend* pada *tweet* yang mereka posting. Oleh karena itu, penelitian ini akan memberikan solusi dalam melakukan identifikasi topik *tweet* yang diposting oleh para Tokoh Publik Politik di Indonesia. *Topic modelling* merupakan model statistik untuk menentukan inti dari topik yang dibahas pada sebuah dokumen. Penggunaan identifikasi *topic modelling* untuk mengetahui topik yang sering dibahas dan menjadi *trend* topik yang menjadi perhatian masyarakat saat ini, sehingga menghasilkan informasi yang lebih ringkas dan akurat untuk kebutuhan informasi masyarakat.

Penelitian ini menggunakan metode *Latent Dirichlet Allocation* (LDA). *Latent Dirichlet Allocation* (LDA) adalah salah satu algoritma untuk *topic modelling* dengan model probabilitas generatif pada koleksi dokumen. Tujuannya agar pemrosesan dokumen pada koleksi data berukuran besar lebih efisien [3]. Metode LDA merupakan salah satu metode *topic modelling* yang populer saat ini, karena mampu memproses data yang berukuran besar. Sehingga pada penelitian *Identifikasi Topik Tweet Tokoh*

Publik Politik Menggunakan Metode Latent Dirichlet Allocation menggunakan metode LDA.

1.2 Rumusan masalah

Perumusan masalah dalam penelitian ini memiliki beberapa masalah yang ingin diteliti sebagai berikut :

1. Berapa banyak jumlah topik yang optimal dalam pemodelan topik dari *tweet* tokoh publik politik menggunakan algoritma *Latent Dirichlet Allocation*?
2. Bagaimana implementasi pemodelan topik *tweet* tokoh publik politik menggunakan algoritma *Latent Dirichlet Allocation*?

1.3 Batasan Masalah

Batasan masalah dalam penelitian ini adalah:

1. Data yang digunakan berupa beberapa *tweet* dari publik politik yang diperoleh melalui *crawling*.
2. Dataset yang digunakan berupa file dokumen dengan ekstensi *csv* yang diperoleh melalui *crawling*.
3. Dataset yang digunakan berupa *tweet* dalam kurun waktu tertentu.
4. Bahasa pemrograman untuk *topic modelling* menggunakan Bahasa pemrograman *python*.
5. Metode yang digunakan untuk *topic-modelling* menggunakan metode *Latent Dirichlet Allocation*.

1.4 Tujuan Penelitian

Penelitian ini memiliki tujuan yang ingin dicapai sebagai berikut:

1. Mengetahui banyak jumlah topik yang optimal dalam pemodelan topik dari *tweet* tokoh publik politik menggunakan algoritma *Latent Dirichlet Allocation*.

2. Mengetahui tentang langkah implementasi dalam pemodelan topik dari *tweet* tokoh publik politik menggunakan algoritma *Latent Dirichlet Allocation*.

1.5 Metode Penelitian

Dalam melakukan pengumpulan data untuk mengidentifikasi permasalahan yang ada, beberapa metode pendekatan yang digunakan antara lain:

1.5.1 Metode Pengumpulan Data

Untuk mendapatkan data yang dibutuhkan dalam pembuatan laporan penelitian digunakan metode sebagai berikut:

1. Studi Pustaka

Merupakan metode pencarian data yang berasal dari buku, jurnal, internet, dan sebagainya yang berhubungan serta mendukung dalam proses pembuatan laporan penelitian.

2. Observasi

Penelitian ini menggunakan metode observasi dengan cara pengamatan secara langsung terhadap semua proses yang dibutuhkan dalam penelitian.

1.6 Sistematika Penulisan

Sistematika penulisan dalam laporan penelitian ini terdiri dari beberapa bagian sebagai berikut :

1. **BAB I : PENDAHULUAN**

Bab ini membahas latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

2. **BAB II : LANDASAN TEORI**

Bab ini menguraikan tentang teori-teori yang digunakan dalam mendukung penelitian, yang menjadi dasar untuk memecahkan masalah dan sebagai studi pustaka yang dipakai dalam penelitian.

3. **BAB III : ANALISIS DAN PERANCANGAN**

Bab ini akan membahas metode yang digunakan, arsitektur umum, langkah preprocessing yang dilakukan, penerapan algoritma *Latent Dirichlet Allocation*, kebutuhan perangkat lunak, dan perancangan aplikasi untuk melakukan identifikasi topik *tweet* tokoh publik politik.

4. **BAB IV : IMPLEMENTASI DAN PEMBAHASAN**

Bab ini berisi tentang pembahasan implementasi dari perancangan sistem, serta hasil pengujian yang diperoleh dari implementasi yang dijabarkan pada bab 3.

5. **BAB V : PENUTUP**

Bab ini merupakan bagian akhir dari laporan penelitian yang berisikan kesimpulan dan saran.

6. **DAFTAR PUSTAKA**

Bab ini berisikan acuan referensi penulisan laporan penelitian yang berasal dari berbagai sumber.

