

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pada awal tahun 2020 masyarakat di media sosial ramai memperbincangkan tentang kebijakan pemerintah Indonesia dalam pengesahan UU Cipta Kerja. Dilansir dari *Kumparan.com*, banyak pengguna media sosial mengungkapkan berbagai komentar mengenai berita hasil dari rapat kebijakan pemerintah dalam pengesahan rancangan undang-undang cipta kerja tersebut. Pendapat atau opini dari *Twitter* dan *Instagram* akan digunakan untuk proses analisis sentimen agar diperoleh informasi yang lebih akurat kemudian dapat diproses untuk membantu pengambilan sebuah keputusan yang tepat dalam menyelesaikan sebuah masalah.

Untuk mengetahui akurasi dari informasi pada dua media sosial yang berbeda, maka dapat dilakukan dengan cara analisis sentimen. Analisis sentimen adalah salah satu cabang penelitian dari *text mining* yang berguna untuk mengklasifikasi dokumen teks berupa opini berdasarkan sentimen [1]. Analisis sentimen bertujuan mengumpulkan polaritas dari sebuah opini yang bersifat positif atau negatif. Media sosial *Twitter* dan *Instagram* dipilih karena keduanya merupakan media sosial yang populer di kalangan masyarakat saat ini. Sebelum dapat melakukan proses analisis sentimen, maka diperlukan pengumpulan data dari kedua media sosial tersebut. Data tersebut dapat diperoleh dengan cara *crawling data*. Data yang telah diperoleh akan diklasifikasi dengan memberi label positif, negatif, atau netral.

Proses pengumpulan data bisa saja mencapai ratusan bahkan ribuan, maka dari itu dibutuhkan sebuah mesin yang dapat menganalisis *tweet* atau *caption* secara otomatis dengan menerapkan salah satu fungsi dari *text mining*, yaitu klasifikasi *text*. Tahap selanjutnya dari analisis sentimen, yaitu *text preprocessing*. *Text preprocessing* dilakukan untuk mengurangi jumlah karakter, mengurangi dimensi proses *preprocessing* dan menghapus karakter yang tidak relevan dengan menggunakan proses *text mining* (*tokenize*, *remove stopword* dan

*stemming*) [2]. Hasil dari *text preprocessing* adalah sebuah dataset yang informatif dan efektif sehingga dapat dianalisis dengan mudah [3].

Penelitian pengklasifikasian analisis sentimen ini dilakukan dengan tujuan mengetahui dan membandingkan respon masyarakat Indonesia di media sosial *Twitter* dan *Instagram* terhadap Undang - undang cipta kerja. Selama ini kita hanya membaca saja perdebatan di media sosial tanpa tahu sejauh mana masyarakat memahami apa yang sebetulnya dikerjakan oleh pemerintahnya. Perlu diketahui, setiap kali pemerintah mengeluarkan kebijakan selalu menimbulkan perdebatan di media sosial. Maka dari itu, opini yang beredar di kedua media sosial *Twitter* dan *Instagram* baik itu positif atau negatif akan dianalisis dan dihitung dengan algoritma yang biasa digunakan untuk proses analisis sentimen yaitu *Naïve Bayes Classifier*.

Undang - undang cipta kerja merupakan salah satu kebijakan pemerintah Indonesia yang cocok untuk diteliti mengingat pada saat pengesahannya memancing banyak perdebatan di media sosial bahkan sampai menyebabkan terjadinya aksi demonstrasi besar oleh buruh dan mahasiswa. Padahal undang - undang ini dibuat agar dapat menyelesaikan masalah yang sering terjadi tumpang tindih dalam dunia kerja di Indonesia. Pemerintah bertujuan supaya dengan adanya uu cipta kerja yang baru dapat memberikan efek positif bagi kedua belah pihak, yaitu bagi pekerja dan perusahaan. Oleh karena itu kebijakan pemerintah tersebut dapat diteliti dan diukur sejauh mana kepentingan dari undang - undang tersebut.

Analisis sentimen dapat mengklasifikasikan opini-opini tersebut dengan menggunakan algoritma *Naïve Bayes Classifier* dimana akan dilakukan perbandingan hasil terhadap media sosial *Twitter* dan *Instagram*. Hasil dari penelitian ini yaitu untuk mengetahui nilai akurasi dan untuk mengetahui respon mengenai setuju atau tidaknya masyarakat terhadap pengesahan UU Cipta Kerja di media sosial yang berbeda yaitu *Twitter* dan *Instagram*. Diharapkan dari penelitian analisis sentimen ini menghasilkan suatu informasi yang berkualitas dengan melihat nilai rasio dari dua media sosial yang berbeda dengan algoritma

*Naive Bayes Classifier* sehingga mendapatkan pandangan terhadap sebuah masalah.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, dapat dirumuskan permasalahannya yaitu bagaimana algoritma *Naive Bayes Classifier* dapat diterapkan untuk mengetahui bagaimana pendapat masyarakat terhadap UU Cipta Kerja di media sosial *Twitter* dan *Instagram*?

## 1.3 Batasan Masalah

Dalam penelitian ini akan diberikan batasan-batasan masalah untuk mengarahkan penelitian ini sesuai dengan tujuannya. Berikut batasan masalahnya:

1. *Tweet* dan *caption* yang diambil menggunakan Bahasa Indonesia
2. *Tweet* dan *caption* yang diambil berupa opini masyarakat tentang UU Cipta Kerja atau #UUCiptaKerja
3. Dalam penelitian ini menggunakan data yang diambil pada bulan Desember 2020 sampai dengan April 2021.
4. *Tweet* dan *caption* yang diambil berupa *text*
5. Algoritma yang digunakan untuk mengukur akurasi dari hasil pengolahan data pada penelitian ini adalah *Naive Bayes Classifier*.

## 1.4 Maksud dan Tujuan Penelitian

Tujuan dari penelitian ini yaitu mengetahui bagaimana alur dan proses untuk mendapatkan data di *Twitter* dan *Instagram* hingga diolah untuk

mendapatkan nilai akurasi yang terbaik dengan menggunakan algoritma *Naïve Bayes Classifier*. Kemudian melakukan perbandingan dan analisis terhadap hasil yang diperoleh.

### **1.5 Manfaat Penelitian**

Manfaat dari penelitian ini adalah dapat memberikan wawasan terhadap apa yang sering diperdebatkan di media sosial.

### **1.6 Metode Penelitian**

Peneliti menjabarkan cara-cara memperoleh data-data yang digunakan untuk kebutuhan penelitian.

#### **1.6.1 Studi Literatur**

Penulis melakukan studi literatur untuk mengumpulkan data teoritis dari berbagai sumber referensi seperti internet, buku-buku dan jurnal ilmiah yang berhubungan dengan judul penelitian.

#### **1.6.2 Analisis Permasalahan**

Analisis permasalahan yang dilakukan pada tahap ini berdasarkan informasi sebelumnya untuk mendapatkan metode yang tepat untuk mengukur keakuratan sebuah data dari dua media sosial yang berbeda.

#### **1.6.3 Implementasi**

Pada tahap ini akan dilakukan pengujian terhadap data dari hasil *preprocessing* dengan algoritma *Naïve Bayes Classifier* sehingga data yang dihasilkan dapat diketahui hasil keakuratan data serta dapat digunakan dalam melakukan perbandingan.

#### 1.6.4 Pengujian

Pada tahap ini dilakukan pengujian pada algoritma *Naïve Bayes Classifier* dengan membandingkan data berdasarkan jumlah opini positif, opini negatif dan netral dari media sosial *Twitter* dan *Instagram*.

#### 1.6.5 Penyusunan Laporan

Tahap akhir yaitu dilakukan penulisan laporan dari keseluruhan penelitian yang telah dilakukan.

### 1.7 Sistematika Penulisan

Secara keseluruhan laporan skripsi ini terdiri dari lima (5) bab sesuai dengan petunjuk laporan skripsi yang berlaku di Universitas Amikom Yogyakarta, sistematika penulisan laporan ini adalah sebagai berikut :

#### BAB I PENDAHULUAN

Berisi latar belakang, rumusan masalah, batasan masalah, maksud dan tujuan penelitian, manfaat penelitian, metode penelitian, dan sistematika penulisan.

#### BAB II LANDASAN TEORI

Pada bab ini akan berisi tentang uraian teori-teori yang digunakan penulis dalam penelitian yang meliputi analisis sentimen, algoritma *Naïve Bayes Classifier*, alur analisis sentimen, dan sosial media yang digunakan dalam penelitian ini

#### BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini berisi tentang analisis dari metode yang digunakan dan juga penerapannya dalam penelitian ini.

#### BAB IV IMPLEMENTASI DAN PEMBAHASAN

Bab ini berisi implementasi dan pembahasan dari apa yang telah dianalisis pada bab sebelumnya serta pengujian hasil yang didapatkan apakah sesuai dengan apa yang diharapkan.

#### BAB V PENUTUP

Bab penutup ini berisikan kesimpulan yang didapatkan penulis melalui bab-bab sebelumnya dan juga menjawab rumusan masalah pada bab 1, serta saran untuk penelitian selanjutnya.

#### DAFTAR PUSTAKA

Bagian ini berisikan daftar referensi – referensi yang telah digunakan dalam penulisan.

