

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Jaminan sosial adalah salah satu bentuk perlindungan sosial yang bertujuan untuk menjamin seluruh rakyat agar dapat memenuhi kebutuhan dasar hidupnya yang layak. Dalam menyelenggarakan jaminan sosial, pemerintah Indonesia membentuk suatu badan hukum bernama Badan Penyelenggara Jaminan Sosial (BPJS) yang diatur pada Undang-Undang Republik Indonesia No. 24 Tahun 2011. Bentuk dari ruang lingkup BPJS dibagi menjadi dua dengan fungsi dan manfaat yang berbeda, yaitu BPJS Kesehatan, dan BPJS Ketenagakerjaan. BPJS Kesehatan adalah ruang lingkup BPJS yang memiliki fungsi utama untuk menyelenggarakan program jaminan kesehatan. BPJS Kesehatan bertugas menyelenggarakan program jaminan kesehatan yang berdasarkan prinsip asuransi sosial dan ekuitas dengan tujuan memberikan manfaat pemeliharaan kesehatan dan perlindungan bagi peserta dalam memenuhi kebutuhan dasar kesehatan sesuai dengan Undang-Undang BPJS dan UU SJSN (Solechan, 2019). BPJS Ketenagakerjaan adalah ruang lingkup BPJS yang memiliki visi untuk mewujudkan jaminan sosial ketenagakerjaan yang terpercaya, berkelanjutan, dan menyejahterakan seluruh pekerja Indonesia. BPJS Ketenagakerjaan berfungsi untuk memperhatikan hak para pekerja dengan 4 program sesuai UU BPJS yaitu program jaminan kecelakaan kerja, jaminan hari tua, jaminan pensiun, dan jaminan kematian (Solechan, 2019)

Banyaknya manfaat yang ditawarkan BPJS menarik banyak peminat dari masyarakat untuk mendaftar menjadi peserta. Hal ini dibuktikan dengan data jumlah peserta program JKN BPJS Kesehatan yang telah mencapai angka 246.679.996 peserta per 31 Januari 2023. Dengan meningkatnya jumlah peserta BPJS, pelayanan dari badan hukum ini pun turut menjadi sorotan masyarakat dan mengakibatkan opini yang tersebar terkait pelayanannya pun bervariasi. Terdapat

beberapa pihak yang pro karena biaya yang diperlukan tidak memberatkan, dan beberapa pihak kontra karena merasa tidak diprioritaskan (Puspita & Widodo, 2021). Opini-opini dari pengguna BPJS dapat menjadi sebuah masukan yang membangun bagi pemerintah untuk meningkatkan kualitas pelayanan sesuai dengan yang dibutuhkan oleh peserta.

Media sosial saat ini termasuk sarana paling populer dalam memperoleh dan membagikan informasi. Banyaknya informasi dalam sosial media membuat beberapa pihak berusaha untuk memahami pendapat pengguna untuk berbagai tujuan seperti reaksi konsumen terhadap promosi suatu produk (Messaoudi et al., 2022). Media sosial yang populer di Indonesia saat ini antara lain Facebook, Youtube, Instagram, Twitter, dan Pinterest.

Media sosial Twitter memfasilitasi penyampaian opini publik di ruang digital. Tweet yang tersedia secara publik menunjukkan reaksi dan engagement suatu pengguna melalui *retweet*, *like*, *mention*, ataupun balasan kepada tweet pengguna yang lain (Karami et al., 2020). Media yang efektif dalam dialektika digital membuat opini terkait diskusi atau produk menjadi dinamis dan konstruktif. Namun di sisi lain, melimpahnya percakapan menyebabkan problematika pada pemahaman percakapan atau arah opini yang lebih dikenal dengan istilah sentimen. Sosial media mengandung banyak informasi berupa opini yang seringkali menyebabkan misinformasi dan penyebaran informasi yang tidak benar (hoax) sehingga sentimen yang ingin disampaikan tidak benar-benar dapat dipahami (Dewi et al., 2022).

Analisis sentimen adalah metode yang digunakan untuk melakukan ekstraksi sentimen dalam suatu teks. Ekstraksi informasi melalui sentimen sudah seringkali dilakukan dengan menggunakan media sosial ataupun website sebagai media dalam pengumpulan datanya. Penelitian terkait analisis sentimen telah dilakukan terhadap beberapa studi kasus seperti ulasan produk, ulasan hotel, hingga topik politik seperti pemilihan presiden. Beberapa metode dalam analisis sentimen yang umum digunakan antara lain Naïve Bayes, K-Means, Decision Tree, K-Nearest Neighbour, Convolutional Neural Network, N-Gram, dan Support Vector Machine.

Pendekatan dalam meningkatkan akurasi analisis sentimen dengan menggabungkan beberapa metode pernah dilakukan oleh Tifani (2020), Kurniawan et al. (2019), dan Hariguna et al. (2019). Pendekatan metode gabungan pada penelitian-penelitian sebelumnya menggabungkan dua metode sebagai bentuk optimasi dalam peningkatan akurasi. Metode-metode yang digabungkan dalam penelitian-penelitian ini antara lain K-Means dengan Naive Bayes, dan N-Gram dengan Naive Bayes. Penggabungan metode ini telah memberikan hasil akurasi yang cukup baik namun kurang optimal. Masih terdapat beberapa kekurangan seperti kesalahan dalam pengelompokan ataupun tingkat akurasi yang masih dapat ditingkatkan lagi.

Berdasarkan hal-hal yang dipaparkan sebelumnya, penelitian ini menawarkan peningkatan akurasi dalam analisis sentimen berupa menggabungkan metode N-Gram, Multinomial Naïve Bayes, dan K-Means. Metode N-Gram digunakan karena metode ini mempertimbangkan konteks dan urutan kata dalam ekstraksi fitur sehingga dapat memberikan informasi yang lebih kaya dari sebuah kalimat. Metode K-Means digunakan untuk clustering sentimen karena algoritma ini membagi data berdasarkan kemiripan data sehingga dapat membantu dalam memahami pola dalam data. Multinomial Naive Bayes digunakan dalam klasifikasi sentimen karena akurasi yang lebih baik jika dibandingkan dengan metode naive bayes yang lain dalam klasifikasi teks.

Penggunaan metode K-Means dan Naïve Bayes dalam penelitian yang dilakukan Hariguna et al. (2019) memberikan kekurangan pada *clustering* data berupa kesalahan klasterisasi data positif menjadi negatif. Untuk membantu dalam lebih mengenali pola data yang digunakan dalam analisis sentiment, dilakukan ekstraksi fitur dengan metode n-gram sebelum memasuki *clustering* data latih. Pengujian penggunaan metode N-Gram sebagai ekstraksi fitur pada teks tweet sebelumnya pernah dilakukan dengan berbagai algoritma *machine learning* dan memberikan hasil yang cukup baik (Gamal et al., 2019). Oleh karena itu, dengan penggabungan ketiga metode yang disebutkan tersebut diasumsikan dapat meningkatkan tingkat akurasi dalam klasifikasi sentimen data tweet menjadi sentimen positif, netral, dan negatif. Metode n-gram dapat membantu dalam clustering data dengan metode k-means dengan mereduksi kata-kata yang tidak

penting, dan mengoptimalkan representasi kata-kata yang akan di kelompokkan. Dengan itu, diasumsikan jika nilai akurasi pengelompokan dengan k-means meningkat maka akurasi multinomial naive bayes akan turut meningkat dalam mengklasifikasikan data sentimen. Pada akhir penelitian ini akan dilakukan pengujian akurasi dari pendekatan tersebut untuk mengetahui tingkat akurasi dari hasil kerja penggabungan ketiga metode tersebut dalam mengklasifikasikan sentimen “BPJS” pada media sosial Twitter.

## **1.2. Rumusan Masalah**

Berdasarkan latar belakang, maka dapat dirumuskan masalah penelitian

1. Bagaimana arah opini masyarakat terhadap BPJS?
2. Bagaimana cara metode N-gram, K-Means, dan Multinomial Naive Bayes dalam mengklasifikasikan opini masyarakat terkait BPJS dalam sentimen positif, netral, dan negatif?
3. Bagaimana tingkat akurasi analisis sentimen dengan metode N-gram, K-Means, dan Multinomial Naive Bayes dalam mengklasifikasikan sentimen opini masyarakat terkait BPJS?

## **1.3. Batasan Masalah**

Adapun Batasan-batasan masalah dalam penelitian yang ditetapkan agar penelitian tidak menyimpang dari tujuan yang sebenarnya:

1. Pengambilan data melalui scraping website Twitter dalam Bahasa Indonesia dengan kata kunci “BPJS”
2. Pengambilan data berjumlah sebesar 500 data tweet
3. Klasifikasi sentimen tweet dilakukan pada 3 kelas yaitu positif, negatif, dan netral.
4. Penelitian ini berfokus pada penilaian akurasi analisis sentimen menggunakan gabungan metode dari N-Gram, Multinomial Naive Bayes, dan K-Means

## **1.4. Tujuan**

Tujuan dari penelitian ini dapat dipaparkan sebagai berikut:

1. Mengetahui arah opini masyarakat terhadap BPJS
2. Mengetahui bagaimana performa metode N-gram, K-Means, dan Multinomial Naive Bayes dalam analisis sentimen tweet BPJS
3. Mengetahui tingkat akurasi dari analisis sentiment dengan metode N-gram, K-Means, dan Multinomial Naive Bayes dalam mengklasifikasikan sentimen tweet BPJS

### **1.5. Manfaat Penelitian**

Dengan dilakukannya penelitian ini, manfaat yang akan didapatkan adalah mengetahui tingkat akurasi analisis sentimen menggunakan gabungan dari metode N-gram, K-Means, dan Multinomial Naive Bayes dalam mengklasifikasikan sentimen data tweet BPJS menjadi 3 kelas yaitu positif, negatif, dan netral. Adapun manfaat lain yang diperoleh dari penelitian ini yaitu mengetahui pandangan masyarakat melalui opini-opini yang tersebar pada media sosial Twitter terkait suatu topik sebagai acuan dalam peningkatan dan perbaikan layanan atau kualitas.