

BAB I

PENDAHULUAN

Bab 1 pendahuluan dijelaskan terkait latar belakang penelitian ini dilakukan. Secara garis besar penelitian ini melakukan *sentiment analysis* media sosial Twitter pada kasus Jaminan Hari Tua menggunakan metode *naive bayes*. Dalam bab ini dijelaskan juga terkait rumusan masalah, tujuan dan manfaat dari penelitian ini

1.1 Latar belakang

Saat ini media sosial sudah sangat berkembang. banyak orang yang mengutarakan opininya di media sosial tentang suatu hal. Media sosial sudah sangat umum digunakan, seperti di Indonesia sendiri biasanya media sosial digunakan untuk sarana mengekspresikan pendapat atau juga sebuah ide akan suatu hal. Data tersebut bisa melalui proses pengekstrakan, pemahaman dan juga pemrosesan data tekstual secara otomatis untuk mendapat sentimen informasi yang terkandung dalam sebuah opini. Contohnya pada media sosial Twitter banyak orang yang mengutarakan pendapat atau opini mereka (Fitri, Andreswari dan Hasibuan, 2019).

Kasus yang digunakan pada penelitian kali ini adalah jaminan hari tua, jaminan kehilangan pekerjaan dan Badan Penyelenggara Jaminan Sosial. Pada saat pekerja akan memasuki usia pensiun, pekerja akan mendapat hak gaji jaminan sosial dan tunjangan sebagai perlindungan sosial. Program Jaminan Hari Tua (JHT) memiliki manfaat yang sangat besar untuk pekerja yang memasuki usia tua atau pensiun (Wijayanti dan Jannah, 2019). Hal itu juga menimbulkan banyak opini yang berisi pro dan kontra terutama di media sosial seperti Twitter (Agara, Santoso dan Puspitawati, 2021).

Oleh karena itu, peneliti melakukan penelitian terkait analisis sentimen media sosial Twitter dengan menggunakan kasus jaminan hari tua. Opini yang disampaikan oleh masyarakat tentang suatu topik pada media sosial tentunya dapat memiliki sentimen yang bersifat negatif atau positif. *Natural Language Processing* atau NLP digunakan untuk mengklasifikasikan sentiment jaminan hari tua. Algoritma *naive bayes* sudah banyak digunakan untuk *text mining* karena memiliki

algoritma yang sederhana namun memiliki akurasi yang cukup tinggi. *Naïve bayes* juga biasa digunakan untuk mengklasifikasikan teks karena memiliki konsep dasar penggabungan probabilitas kata dan kategori dari dokumen (Novendri *et al.*, 2020).

1.2 Perumusan Masalah

Adapun rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana hasil analisis sentimen pengguna Twitter mengenai Jaminan Hari Tua menggunakan metode *Naïve Bayes*?
2. Bagaimana akurasi analisis sentimen metode *naïve bayes* menggunakan ekstraksi fitur *Bag of Word*?

1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Mengetahui hasil analisis sentimen pengguna Twitter mengenai Jaminan Hari Tua menggunakan metode *Naïve Bayes*
2. Mengetahui akurasi analisis sentimen metode *naïve bayes* menggunakan ekstraksi fitur *Bag of Word*

1.4 Batasan Penelitian

Adapun batasan masalah pada penelitian ini adalah sebagai berikut:

1. Data yang digunakan berasal dari media sosial Twitter dengan kata kunci jaminan hari tua yang berbahasa Indonesia menggunakan.
2. Metode yang digunakan untuk analisis sentimen adalah *Naïve Bayes*
3. Bahasa pemrograman yang digunakan adalah *Python* dengan menggunakan *tools Google Collaboraty*

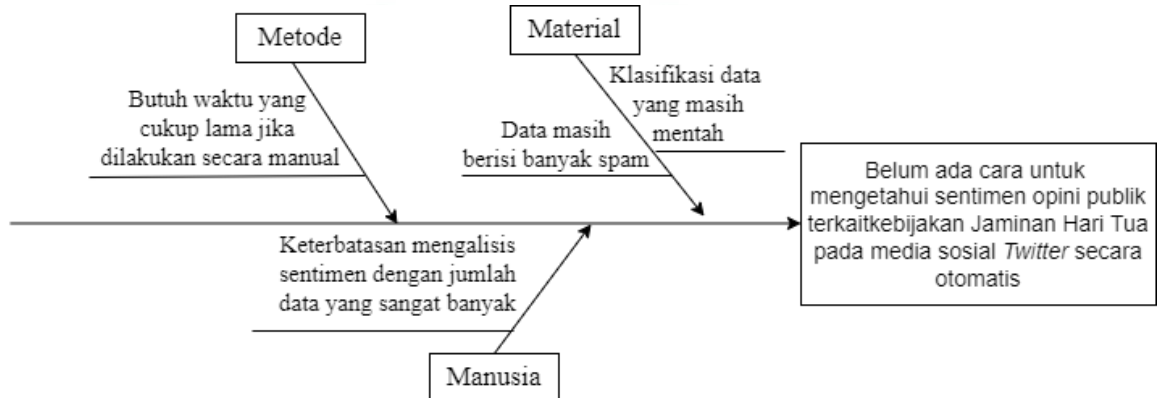
1.5 Manfaat Penelitian

Adapun manfaat pada penelitian ini adalah sebagai berikut:

1. Berisi pengetahuan mengenai analisis sentimen pengguna Twitter terkait Jaminan Hari Tua menggunakan metode *Naïve Baye*
2. Dapat mengetahui opini pengguna Twitter terkait program Jaminan Hari Tua lebih mengarah positif atau negatif

1.6 Kerangka Pemikiran

Kerangka pemikiran pada penelitian ini dimodelkan menggunakan diagram sebab akibat atau *fishbone diagram*. Diagram ini menjelaskan alur berpikir dalam pengerjaan penelitian berdasarkan masalah hingga penyelesaian yang diusulkan.



Gambar 1.1 Diagram *Fishbone Sentiment Analysis*

Pada Gambar 1.1 terdapat diagram *fishbone* menjelaskan beberapa hal yang menjadi kategori permasalahan. Kategori permasalahan tersebut yang pada akhirnya menjadi latar belakang dari dilakukannya penelitian. Dalam diagram tersebut terdapat 3 kategori permasalahan yaitu Metode, Material dan Manusia. Metode yang digunakan tidak memakan banyak waktu saat melakukan ulasan pada Twitter. Terbaginya *tweet* yang membahas tentang program Jaminan Hari Tua menjadi data yang bisa di *sentiment analysis*. Dari kumpulan data yang tidak stabil tersebut maka dapat dibagi menjadi *tweet* yang bersifat positif dan negatif yang berguna untuk pemerintah maupun masyarakat. *Tweet* yang digunakan menjadi data adalah *tweet* dari pengguna Twitter yang berbahasa Indonesia. Lalu untuk kategori material, banyaknya data yang berupa *spam* yang berisi *link* dan sampah seperti emotikon serta karakter tidak penting yang tidak memiliki nilai. Dan untuk kategori manusia, rentan terjadinya *human error* jika dilakukan secara manual dan memiliki data yang sangat banyak.