

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Menurut Cahyono (2017), pertumbuhan pada media sosial secara online seperti pada twitter dapat memunculkan informasi tekstual yang masih harus digali nilai informasi pada twitter. Informasi tekstual dikategorikan antara fakta dan opini. Fakta adalah ekspresi objektif, sedangkan opini adalah ekspresi subjektif yang menggambarkan sentimen seseorang, pendapat tentang sebuah kejadian. Pada media sosial twitter terdapat istilah tweet yang memiliki arti suatu status atau yang dibuat oleh penulis atau penggunanya. Tweet dapat memberikan suatu ekspresi berupa perasaan atau keadaan dari pengguna Twitter. Menurut Haqqizar & Larasyanti (2019) opini tersebut dapat digunakan dan dimanfaatkan sebagai bahan data pengembangan produk mereka, agar dapat saling bersaing dan bertahan terhadap pesaing lain. Salah satunya perusahaan di bidang telekomunikasi yaitu provider dalam hal internet.

Menurut Wibisono (2016) kebutuhan internet dibutuhkan untuk dapat mengakses informasi di era saat ini. Hal tersebut dikarenakan dengan internet masyarakat dapat menjelajahi dunia internet untuk mencari informasi yang dibutuhkan, sehingga provider menjadi jawaban dari kebutuhan tersebut yaitu sebagai suatu wadah atau perusahaan yang menyediakan layanan jasa dalam mengakses internet. Menurut Fransiska & Gufroni, (2020) by.U salah satu layanan provider yang baru sebagai penyedia layanan serba digital pertama di Indonesia yang diluncurkan Telkomsel, dan juga terdapat MPWR yang diluncurkan Indosat. By.U dan MPWR merupakan layanan seluler yang menawarkan pemilihan nomor, penetapan kuota, pengiriman sim card serta layanan live chat serba online. Untuk mengetahui preferensi konsumen antara provider by.U dan MPWR dibutuhkan dari opini pelanggan.

Pelanggan sangat dibutuhkan oleh perusahaan dan hal tersebut tergantung pada perusahaan dalam menjalin hubungan dengan pelanggan. Maka dari itu perusahaan provider harus mencari cara untuk meningkatkan pelayanan dengan mencari informasi tentang kepuasan konsumen. Twitter merupakan suatu media

yang dapat menjadi penyebar informasi secara cepat. Informasi yang dapat berupa komentar, opini, kritik maupun saran yang bersifat positif dan negatif (Haqqizar & Larasyanti, 2019). Tweet yang di tulis oleh pelanggan dijadikan data informasi untuk perhitungan opini pelanggan terhadap layanan provider dengan keyword #by.U dan #MPWR pada twitter. Kemudian data tweet dikategorikan berdasarkan sentimen analisis berupa positif serta negatif dan disertai aspek yang sesuai dengan kepuasan konsumen produk provider dan juga opini yang sering dibicarakan pada fitur provider. Menurut Wibisono (2016), aspek aspek pendorong kepuasan diantaranya adalah kualitas produk, harga, layanan dan kemudahan penggunaan atau jaringan. Metode yang digunakan pada sentimen analisis dengan kategori positif dan negatif yaitu metode *Multinomial Bayes* dan metode untuk mengelompokkan tweet menggunakan algoritma *Cosine Similarity* dengan menghitung kemiripannya berdasarkan *Kata kunci* atau kata kunci.

Analisis Sentimen merupakan salah satu dari *Machine Learning* dan *Natural Language Processing (NLP)*. Cara kerjanya adalah dengan melakukan klasifikasi terhadap opini positif dan negatif. Analisis Sentimen berguna dalam menganalisa pendapat orang, sentimen, evaluasi, penilaian, sikap, dan emosi terhadap entitas seperti produk, jasa, masalah, peristiwa atau suatu topic. Metode yang digunakan pada analisis sentimen ada banyak beberapa diantaranya Naïve Bayes, SVM, Neural Network dan KNN (Cahyono, 2017). Sedangkan, *Cosine Similarity* merupakan suatu metode menghitung tingkat similaritas atau kemiripan antar dua objek atau lebih. Perhitungan ini terdiri dua buah objek *tweet* komentar ( $D1$  dan  $D2$ ) untuk menghitung sebuah *similarity* antar *tweet* komentar yang dinyatakan dalam sebuah *vector* dengan menggunakan *keyword* atau kata kunci (Sya'bani & Umilasari, 2018). Sehingga dengan kedua metode ini dapat berguna bagi produsen sebagai informasi mengenai kategori produk yang tepat untuk ditingkatkan dan bagi konsumen berguna untuk memilih provider secara detail dengan menyesuaikan kebutuhan.

Penelitian terkait sudah pernah dilakukan oleh Haqqizar N & Larasyati T.K, (2019) mengenai “*analisis sentiment* terhadap layanan telekomunikasi telkomsel di twitter dengan metode *Naïve Bayes*”, penelitian tersebut menggunakan 151

data tweet dan berisi tentang mengklasifikasi *tweet* opini pada layanan telkomsel tanpa ada klasifikasi fitur dengan menggunakan metode *Naïve Bayes* dan memperoleh nilai akurasi tertinggi 70.21%, *precision* 70.11%, dan *recall* 70.33%. Sedangkan pada penelitian Rizki Tri Wahyuni, Dhidik Prastiyanto, dan Eko Suprpto (2017) yang berjudul “Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi” mengklasifikasikan dokumen secara otomatis ke dalam folder berbeda pada database agar lebih mengelola dokumen yang ada dan tingkat ketepatan senilai 98% dari data berjumlah 50 dokumen skripsi dan terdiri 9 kategori skripsi. Dari sumber penelitian yang didapat, maka dari itu akan digabungkan antara penerapan algoritma *Cosine Similarity* sebagai klasifikasi tweet dan metode *Multinomial Naïve Bayes* sebagai metode analisis sentimen pada kasus preferensi konsumen pada produk by.U dan MPWR di media twitter serta perhitungan akurasi, *precision* dan *recall* dikarenakan metode ini cocok dalam pengklasifikasian teks atau dokumen. Penelitian ini berguna bagi produsen sebagai informasi fitur apa yang harus ditingkatkan pelayanannya dan sebagai konsumen dapat membantu untuk memilih provider sesuai kebutuhan.

Berdasarkan latar belakang yang dijelaskan tentang preferensi konsumen terhadap produk by.U dan MPWR, maka penulis akan membuat penelitian dengan judul “Preferensi Konsumen terhadap Produk by.U dan MPWR dengan Analisis Sentimen Berbasis *Multinomial Naïve Bayes*”.

## 1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, maka dapat dirumuskan rumusan masalah yang akan di bahas dalam penelitian ini yaitu:

1. Berapa jumlah tweet pada setiap aspek kategori fitur di media twitter pada provider by.U dan MPWR ?
2. Berapa persentase tingkat akurasi, *recall* dan *precision* yang diperoleh dari perhitungan metode *Multinomial Naïve Bayes* pada masing masing dataset provider by.U dan MPWR ?

### 1.3. Tujuan

Berdasarkan latar belakang di atas, tujuan pada penelitian ini yaitu:

1. Untuk dapat mengetahui jumlah tweet pada setiap parameter aspek kategori fitur di media twitter pada provider by.U dan MPWR
2. Untuk dapat mengetahui nilai akurasi, *recall* dan *precision* pada perhitungan dengan menggunakan metode *Multinomial Naïve Bayes* pada masing masing dataset provider by.U dan MPWR.

### 1.4. Manfaat

Hasil penelitian ini dapat diharapkan bermanfaat bagi peneliti yang lain tentang menghitung nilai akurasi, *recall* dan *precision* menggunakan metode algoritma *Multinomial Naive Bayes* dan preferensi konsumen terhadap layanan provider by.U dan MPWR pada media sosial twitter.

### 1.5. Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Data yang telah digunakan menggunakan data twitter di bulan Desember.
2. Data yang telah digunakan adalah postingan berbahasa Indonesia
3. Data diambil menggunakan API (*Application Programming Interface*) Twitter dengan API Key penulis.
4. Pengklasifikasian berdasarkan 2 kategori yaitu positif dan negatif.
5. Penelitian ini menggunakan pemrograman python dan software Jupyter Notebook untuk melakukan proses analisis data.