

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Teknologi yang berkembang sekarang ini semakin canggih, teknologi yang paling populer saat ini salah satunya adalah internet karena dengan adanya internet banyak informasi dan data yang berharga dapat diambil dengan cepat dan mudah. Penggunaan internet terbanyak di Indonesia pada tahun 2016 digunakan untuk membuka media sosial dengan presentase mencapai 73,30% (kominfo.go.id, 2016). Media Sosial adalah media yang digunakan oleh konsumen untuk berbagi teks, gambar, suara, video, dan informasi baik dengan orang lain maupun perusahaan dan vice versa (Kotler & Keller, 2012).

Salah satu perkembangan teknologi yang memiliki andil besar dalam memberikan kemudahan bagi manusia untuk berkomunikasi masih digunakan sampai sekarang adalah *Twitter*. *Twitter* berdiri sejak tahun 2006 dan telah digunakan diseluruh penjuru dunia. Saat ini sekitar 500 juta lebih ulasan yang dikirimkan tiap harinya dan memiliki lebih dari pengguna aktif lebih dari 307 juta pengguna aktif (Stat, 2017).

Twitter dikenal dengan layanan *micro-blogging*, memungkinkan penggunanya membaca, menulis atau mengirimkan pesan pada forum umum (Setiawan & Setyohadi, 2017). *Twitter* membuat perusahaan dapat

melakukan interaksi dengan pelanggan secara tidak langsung.



Perusahaan dapat memberikan informasi terbaru pada pelanggan, dan mendapatkan timbal balik dari pelanggan dengan cepat, sehingga perusahaan dapat membuat evaluasi mengenai layanan yang diberikan. Demikian juga yang dilakukan oleh perusahaan penyedia layanan telekomunikasi seluler saling bersaing untuk meningkatkan inovasi produk, layanan dan mutu produk mereka.

Penggunaan telekomunikasi seluler di Indonesia pada tahun 2015 mencapai 351.860.784 (Statistik, 2016). Terdapat tiga penyedia layanan telekomunikasi seluler yang paling populer di Indonesia yaitu XL Axiata, Telkomsel, Indosat (Socialbakers, 2017). Hal ini membuat produk-produk *provider* telekomunikasi berlomba-lomba melakukan pemasaran via media sosial. Perusahaan membuat akun media social Twitter dan kemudian menunggu respon dari pengguna atau pelanggan *provider* telekomunikasi untuk mengetahui bagaimana respon atas produk dan layanan yang mereka berikan. Perusahaan dapat mengetahui tingkat kepuasan dan keluhan dari pelanggan, maka diperlukan analisis sentimen pada Twitter pengguna yang berhubungan dengan layanan *provider* telekomunikasi. Analisis sentimen atau opinion mining adalah studi komputasional dari opini-opini orang, sentimen, dan emosi melalui entitas dan atribut yang dimiliki yang diekspresikan dalam bentuk teks (Liu, 2012). Analisis sentimen dapat mengelompokkan polaritas dari teks dalam kalimat atau dokumen tersebut apakah termasuk positif atau negatif (Rofiqoh, Perdana, & Fauzi3, 2017).

Telkomsel merupakan salah satu *provider* telekomunikasi yang memiliki pelanggan terbanyak. Hal ini dapat diketahui dari hasil survey yang dilakukan oleh Databoks 2016 yang menunjukkan hasil bahwa terdapat 157,4 juta pelanggan dari Telkomsel, 85 juta pelanggan dari Indosat dan 44 juta pelanggan dari XL Axiata (Databoks, 2016). Hasil tersebut menunjukkan bahwa banyak pengguna *provider* Telkomsel yang digunakan di Indonesia.

Respon pengguna *Twitter* melalui *tweets* terhadap *provider* telekomunikasi akan dijadikan informasi atau biasa disebut dengan big data. Data yang besar tersebut dapat dilakukan analisis terhadap informasi atau opini-opini yang telah tersimpan. Data-data tersebut dapat diolah dengan menggunakan teknik big data, dan dapat dilakukan analisis dengan menggunakan metode dari data mining, seperti metode *Naïve Bayes Classifier (NBC)*. *Naïve Bayes Classifier* merupakan teknik pembelajaran mesin yang berbasis probabilitas dan metode yang sederhana tetapi memiliki akurasi serta performansi yang tinggi dalam pengklasifikasian teks (Ling, N, Kencana, & Bagu, 2014). Metode *Naïve Bayes* bisa digunakan untuk mengklasifikasikan opini ke dalam opini positif, netral dan negatif. Metode *Naïve Bayes* bisa berfungsi baik dalam proses pengklasifikasian teks dan menyelesaikan masalah *opinion orientation* pada *opinion mining* terutama kasus dataset yang digunakan dalam proses training dan testing NBC (Rozi, Pramono, & Ac, Implementasi Opinion Mining (Analisis, 2012)).

Provider telekomunikasi merupakan layanan yang menjangkau ke seluruh pelosok Indonesia, ada ribuan percakapan per harinya mengenai layanan yang diberikan di *provider* telekomunikasi media sosial, salah satunya melalui *Twitter*. Opini yang diberikan sangat beragam, terdapat opini negatif dan opini positif. Hal ini dilakukan dengan proses analisis dan sangat penting untuk setiap perusahaan *provider* telekomunikasi itu sendiri untuk mengetahui sentimen atau opini masyarakat mengenai layanan yang diberikan, kekurangan dari layanan yang diberikan dan masukan dari masyarakat tentang *provider* itu sendiri, supaya perusahaan dapat memperbaiki layanan yang diberikan kepada masyarakat melalui opini yang diberikan melalui *tweets* yang diambil dari *Twitter*.

1.2. Rumusan Masalah

Berdasarkan penjelasan latar belakang di atas, maka dapat dirumuskan masalah sebagai berikut:

1. Bagaimana menganalisis dan mengevaluasi pendapat masyarakat tentang *provider* telekomunikasi yang ada di Indonesia melalui *tweet* pada *Twitter*?
2. Bagaimana mengklasifikasikan setiap sentimen pada *tweet* yang akan dilakukan analisis?

1.3. Batasan Masalah

Batasan masalah untuk penelitian ini memiliki ruang lingkup sebagai berikut:

1. Data yang digunakan untuk melakukan penelitian ini adalah data dari Twitter .
2. Data yang digunakan dalam penelitian adalah data *tweets* menggunakan Bahasa Indonesia.
3. *Provider* telekomunikasi yang akan di lakukan analisis opini terhadap salah satu *provider* besar di Indonesia yaitu Telkomsel.
4. Lama waktu pengambilan data dan penelitian ini adalah dimulai dari tanggal 01 Oktober 2017 sampai dengan tanggal 31 Januari 2018.
5. Hasil keluaran dari analisis sentimen dengan menggunakan metode *Naïve Bayes Classifier* berupa sentimen positif, negatif dan netral.

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah:

1. Menganalisis dan mengevaluasi sentimen masyarakat tentang *Provider Telekomunikasi Indonesia* terkhusus *Telkomsel* pada *Twitter* menggunakan *Naïve Bayes Classifier*.
2. Mengklasifikasikan sentimen pada *tweets*.

1.5. Metodologi Penelitian

Penelitian ini dilakukan dengan berbagai tahap sehingga dapat menghasilkan hasil yang akurat. Sehingga, metodologi yang dapat digunakan dalam penelitian ini adalah sebagai berikut:

a) Studi Literatur

Pada tahap ini merupakan proses mempelajari dari sumber referensi berupa textbook, artikel ilmiah,

dan jurnal yang berkaitan dengan penelitian analisis sentimen masyarakat tentang *provider* telekomunikasi Indonesia pada twitter menggunakan *Naïve Bayes Classifier*.

b) Pengumpulan Data

Pengumpulan data diambil dari Twitter. Data yang akan digunakan adalah data dari 01 Oktober 2017 sampai 31 Januari 2018. Data yang diambil merupakan data yang berhubungan dengan *provider* telekomunikasi Indonesia.

c) Pre-processing

Tahap pertama setelah data terkumpul dimulai 01 Oktober 2017 sampai 31 Januari 2018 maka akan dilakukan tahap pre-processing, tahap pre-processing dilakukan dengan tahapan sebagai berikut :

1. Seleksi *Tweets*

Pada tahap ini hal yang dilakukan adalah melakukan seleksi dari status yang memiliki hubungan atau mengandung kata-kata kunci dari khusus ini menggunakan kata kunci "Telkomsel".

2. *Stopwords*

Berdasarkan data-data status yang telah didapatkan tidak dapat langsung digunakan sebagai informasi yang valid, dimana data-data tersebut harus dibersihkan dahulu. Seperti proses menghilangkan kata-kata yang tidak memiliki arti

atau menghilangkan kata- kata yang maknanya tidak diperlukan. Kata-kata tersebut seperti yang, dari dan kata penghubung lainnya.

3. Parsing/Tokenisasi

Proses ini merupakan proses dimana memecah dokumen menjadi potongan urutan karakter dan sebuah set dokumen yang diberikan menjadi potongan-potongan kata atau karakter yang sesuai dengan kebutuhan sistem.

4. Stemming

Proses ini merupakan proses dimana data dari status-status yang sudah didapatkan melalui proses stopwords akan dilakukan proses *stemming* yaitu menghilangkan imbuhan yang terdapat dalam kata.

Seperti: Memakai → Pakai

Berdasarkan contoh diatas adalah proses menghilangkan kata imbuhan mem-.

d) Ekstraksi Opini

Pada tahap ini akan dilakukan proses penyaringan *tweets* yang telah dilakukan preprocessing data sebelumnya untuk mendapatkan tweet yang hanya mengandung opini atau sentimen.

e) Penerapan Algoritma

Pada tahap ini akan dilakukan penggunaan algoritma dengan data yang telah terkumpul untuk didapatkan hasil akhir dari analisis sentimen.

1.6. Sistem Penulisan Tugas Akhir

Sistematika penulisan tugas akhir adalah sebagai berikut:

BAB 1 PENDAHULUAN

BAB 2 TINJAUAN PUSTAKA

BAB 3 LANDASAN TEORI

BAB 4 METODOLOGI PENELITIAN

BAB 5 HASIL DAN PEMBAHASAN

BAB 6 KESIMPULAN DAN SARAN