

BAB I

PENDAHULUAN

1.1 Latar Belakang

Transportasi merupakan hal yang erat hubungannya dengan masyarakat. Transportasi terdapat berbagai macam jenisnya, baik roda dua maupun roda empat. Transportasi umum juga masih digunakan masyarakat sebagai sarana untuk menuju suatu tempat walaupun untuk memiliki kendaraan pribadi di zaman sekarang sangatlah mudah (Nugroho, et al 2016). Di Indonesia transportasi umum masih menjadi pilihan oleh masyarakat, beberapa tahun terakhir hingga saat ini transportasi umum mengalami perkembangan yang sangat pesat karena adanya transportasi umum yang berbasis *online*. Berkembangnya teknologi saat ini memunculkan aplikasi transportasi berbasis *online* salah satunya yaitu aplikasi gojek. Gojek merupakan transformasi dari ojek konvensional yang bertempat di suatu pangkalan untuk menunggu pelanggan, atau bisa disebut ojek pangkalan. Kemudahan dalam pemesannya dengan menggunakan teknologi *smartphone* membuat gojek semakin banyak diminati, hal ini dapat menjadi solusi transportasi di Indonesia terutama di kota besar yang sering sekali mengalami kemacetan.

Gojek merupakan sebuah perusahaan teknologi asal Indonesia yang menyediakan jasa transportasi ojek, saat ini aplikasi gojek telah diunduh oleh 170 juta pengguna di Indonesia dan Asia Tenggara (cnbcindonesia.com). Gojek Indonesia merupakan satu-satunya perusahaan dari Asia Tenggara yang masuk

dalam daftar 56 perusahaan yang mengubah dunia atau disebut dengan daftar “*Change the World*”, dimana gojek menempati peringkat ke-17, bersama sejumlah perusahaan-perusahaan kelas dunia lainnya (gojek.com). Berbeda dari pesaingnya, gojek merupakan *startup* yang menghadirkan berbagai layanan yang berjumlah 23 layanan, diantaranya GO-RIDE, GO-CAR, GO-FOOD, GO-PAY, GO-SEND, GO-PULSA dan lain sebagainya. Indonesia mengalami peningkatan pengguna internet menjadi 143.26 juta orang , setara dengan 54.7% berdasarkan dari hasil survei APJII pada tahun 2018. Dari hasil survei tersebut dapat diketahui bahwa perkembangan teknologi di Indonesia cukup tinggi dan memiliki potensi untuk dikembangkan dalam berbagai aspek.

Pertumbuhan dalam penggunaan media sosial juga merupakan salah satu pengaruh dari peningkatan jumlah konsumen. Pengguna media sosial yang semakin meningkat setiap tahunnya menjadikan *tools* perusahaan untuk melihat konsumennya yang membuat media sosial digunakan sebagai sarana kegiatan bisnis. *Twitter* merupakan media sosial yang digunakan untuk memudahkan bisnis dalam menampilkan jasa yang mereka tawarkan, sehingga konsumen dapat melihat dan menggunakan jasa tersebut.

Di dunia bisnis, perusahaan melakukan strategi agar jasa yang ditawarkan menarik dan banyak yang menggunakan aplikasi tersebut, salah satu strategi yang dilakukan untuk meningkatkan jumlah pengguna adalah dengan memiliki strategi pemasaran. Strategi yang digunakan oleh gojek kali ini adalah dengan menambahkan promo-promo serta berbagai macam voucher yang bisa digunakan oleh konsumennya. Adanya strategi pemasaran merupakan salah satu usaha yang

dapat meningkatkan jumlah pengguna jasa ojek *online*. Sepanjang tahun 2018 Gojek Indonesia telah berkontribusi Rp 55 triliun ke perekonomian Indonesia (cnbcindonesia.com).

Data yang ada di dalam media sosial disebut *User Generated Content* (UGC). Di UGC ada bagian dalam bentuk opini publik. Data di media sosial yang diposting oleh pengguna memiliki berbagai karakteristik, dan tersedia untuk pengguna lain, dalam data tersebut dapat berisi sejumlah upaya dan pendapat kreatif. Salah satu media sosial yang memiliki UGC adalah *twitter*. *Twitter* adalah teks, foto, video, dan url berbasis media sosial yang dapat digunakan oleh pengguna ke pengguna lain. Di UGC banyak sekali sentimen yang ditulis dan diungkapkan oleh publik, ada bentuk tanggapan positif dan negatif data yang dapat diolah dan digunakan oleh perusahaan, sehingga perusahaan dapat memahami sentimen apa yang mendominasi pemikiran konsumen tentang perusahaan atau produk mereka.

Ulasan dari pengguna sering digunakan sebagai alat yang efektif dan efisien dalam menemukan informasi terhadap suatu produk atau jasa. Menurut Sung 2014 dalam Khotimah 2019, bahwa penelitian baru-baru ini menemukan hampir 50% dari pengguna internet bergantung pada rekomendasi *word of mouth* (opini) sebelum menggunakan suatu jasa ojek *online*, karena *review* dari pengguna lain dapat menyediakan informasi terbaru dari produk atau jasa tersebut berdasarkan perspektif pengguna-pengguna yang sudah menggunakan produk atau jasa tersebut (Sung dan Sung 2014 dalam Khotimah 2019).

Pelanggan yang merasa tidak puas dengan layanan atau jasa yang ditawarkan oleh sebuah perusahaan *startup* biasanya akan menuliskan keluhannya di media sosial. Di sisi lain, ada juga pelanggan yang merasa puas, yang mengekspresikan sikap positif mereka terhadap sebuah layanan atau produk di media sosial. Tanpa disadari, opini-opini pelanggan yang dituliskan di media sosial akan memberikan pengaruh pada calon konsumen. Akan tetapi, memantau opini dari masyarakat juga bukanlah hal yang mudah. Opini yang dimuat di media sosial jumlahnya terlalu banyak. Oleh sebab itulah, diperlukan sebuah metode atau teknik khusus yang mampu mengkategorikan *review-review* tersebut, apakah termasuk positif atau negatif (Potdar, et al 2016 dalam Khotimah 2019), serta topik apa saja yang sering dikeluhkan oleh pengguna media sosial mengenai layanan pada perusahaan Gojek Indonesia.

Analisis sentimen adalah suatu proses dalam mendeteksi teks secara kontekstual, ini menentukan suatu teks memiliki makna positif dan negatif. Teknik ini digunakan di banyak bidang seperti *e-commerce*, perawatan kesehatan, hiburan dan politik. Sebagai contoh, analisis sentimen berguna bagi perusahaan untuk memantau opini konsumen mengenai layanan dan produk mereka, dan bagi konsumen untuk memilih produk terbaik berdasarkan opini publik (Khotimah 2019). Penelitian ini akan menggunakan metode *Naïve Bayes Classifier* dengan visualisasi *Latent Dirichlet Allocation*. Penggunaan *Naïve Bayes Classifier* dalam proses klasifikasi memiliki kelebihan dalam kemudahan penggunaannya serta waktu pemrosesannya yang cepat, mudah diimplementasikan dengan struktur yang cukup sederhana dan tingkat efektifitas yang tinggi. *Latent Dirichlet*

Allocation merupakan sebuah metode yang tergolong sebagai salah satu bentuk dari *text mining* untuk menemukan pola tertentu pada sebuah dokumen. *Latent Dirichlet Allocation* membuat dokumen menghasilkan beberapa macam topik yang berbeda, sehingga tidak secara spesifik mengelompokkan dokumen kedalam sebuah topik tertentu. Hasil dari analisis pemodelan topik menggunakan *Latent Dirichlet Allocation* akan ditampilkan dengan sebuah visualisasi berupa *dashboard* pada media sosial yang dapat menampilkan *trend* dari topik-topik yang sering dibahas, sebagai suatu sistem yang dapat membantu pengambilan keputusan. Oleh karena itu, pada penelitian ini penulis akan melakukan Analisis Sentimen Pada *Twitter* Gojek dengan Metode *Naïve Bayes Classifier* Menggunakan Visualisasi *Latent Dirichlet Allocation*.

Beberapa penelitian mengenai Analisis Sentimen Pada *Twitter* Gojek dengan Metode *Naïve Bayes Classifier* Menggunakan Visualisasi *Latent Dirichlet Allocation* telah dilakukan oleh Agustina, A (2017) dengan judul “Analisis dan Visualisasi Suara Pelanggan Pada Pusat Layanan Pelanggan dengan Pemodelan Topik Menggunakan *Latent Dirichlet Allocation* (LDA) Studi Kasus: PT. Petrokimia Gresik” dengan hasil pengujian yang didapatkan yaitu pengujian dokumen terhadap topik menghasilkan ketepatan sebesar 83.7% yaitu 190 dokumen dari 227 dokumen.

Penelitian yang dilakukan oleh Kamal dan Widjajanto (2017) dengan judul penelitian : “*Text Mining* Untuk Analisa Sentiment Ekspedisi Jasa Pengiriman

Barang Menggunakan Metode *Naïve Bayes* Pada Aplikasi J&T Express” dengan hasil yang didapatkan akurasi sebesar 83%.

Penelitian yang dilakukan oleh Hadi et.al (2017) dengan judul penelitian : “*Text Mining* Pada Media Sosial Twitter Studi Kasus: Masa Tenang Pilkada DKI 2017 Putaran 2” dengan hasil yang didapatkan akurasi rata-rata sebesar 80%.

Penelitian yang dilakukan oleh Syakuro (2017) dengan judul penelitian : “Analisis Sentimen Masyarakat Terhadap *E-Commerce* Pada Media Sosial Menggunakan Metode *Naïve Bayes Classifier* (NBC) dengan Seleksi Fitur *Information Gain* (IG)” dengan hasil yang didapatkan akurasi sebesar 88.8%.

Penelitian yang dilakukan oleh Utami (2017) dengan judul penelitian : “Analisis Topik Data Media Sosial Twitter Menggunakan Model Topik *Latent Dirichlet Allocation*” dengan hasil yang didapatkan pemodelan topik LDA pada kelima lokasi *tweet* di Kota Bogor dan rentang waktu tertentu berhasil membentuk topik dengan informasi atau deskripsi topik untuk setiap lokasi *tweet*.

Penelitian yang dilakukan oleh Prihatini et.al (2017) dengan judul penelitian : “Metode *Latent Dirichlet Allocation* Untuk Ekstraksi Topik Dokumen” dengan hasil yang didapatkan metode LDA memiliki kinerja sangat baik dalam melakukan ekstraksi topik untuk dokumen teks.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, permasalahan yang akan dibahas dalam penelitian ini adalah sebagai berikut :

1. Bagaimana klasifikasi sentimen pada *twitter* gojek menggunakan metode *Naïve Bayes Classifier* dan hasil akurasinya?
2. Bagaimana visualisasi *trend* topik sentimen pada *twitter* gojek menggunakan *Latent Dirichlet Allocation* ?

1.3 Tujuan Penelitian

Adapun tujuan penelitian berdasarkan rumusan masalah tersebut yaitu :

1. Mengklasifikasikan sentimen pada *twitter* gojek menggunakan metode *Naïve Bayes Classifier* dan melihat hasil akurasi.
2. Mengetahui *trend* topik pada sentimen *twitter* gojek menggunakan *Latent Dirichlet Allocation*.

1.4 Manfaat Penelitian

1. Manfaat Teoritis

Membantu perkembangan ilmu pengetahuan mengenai metode *text mining* sehingga dapat digunakan sebagai bahan bacaan dan referensi bagi pembaca dalam melakukan penelitian yang berhubungan dengan *text*

mining terutama pada metode Analisis Sentimen Pada *Twitter* Gojek dengan Metode *Naïve Bayes Classifier* Menggunakan Visualisasi *Latent Dirichlet Allocation*.

2. Manfaat Praktis

- a. Bagi Peneliti, peneliti mampu menerapkan metode yang sesuai dalam materi yang telah dipelajari serta peneliti mempunyai pengetahuan dan wawasan mengenai *text mining* menggunakan Analisis Sentimen Pada *Twitter* Gojek dengan Metode *Naïve Bayes Classifier* Menggunakan Visualisasi *Latent Dirichlet Allocation*.
- b. Bagi pihak terkait diantaranya perusahaan, masyarakat, dan investor yaitu dapat membantu mengetahui sentimen pengguna *twitter* terhadap transportasi *online* Gojek Indonesia sehingga masyarakat terutama perusahaan dapat mengetahui sentimen tentang Gojek, dan memudahkan perusahaan dalam mengambil kebijakan untuk mengatasi sentimen negatif.

1.5 Batasan Masalah

1. Data yang digunakan adalah data *tweet* dari media sosial *twitter* berbahasa Indonesia yang diambil menggunakan *twitter* API periode Januari 2020.
2. Evaluasi klasifikasi menggunakan tingkat akurasi
3. Pada penelitian ini membagi data *testing* dan data *training* menggunakan *40-fold Cross Validation* dengan pembagian data 60% : 40%