

BAB I

PENDAHULUAN

Bab ini berisikan tentang latar belakang, perumusan masalah, batasan masalah, tujuan, manfaat, dan sistematika penulisan yang digunakan dalam laporan skripsi ini.

1.1. Latar Belakang

Berdasarkan data *survey* pertumbuhan pengguna internet di Indonesia mengalami peningkatan setiap tahunnya, menurut hasil survei Asosiasi Jasa Internet Indonesia (APJII), Pengguna internet aktif di Indonesia pada tahun 2018 telah mencapai 64,8% dari total jumlah penduduk sebanyak 264,16 juta jiwa. Data ini berasal dari hasil *survey* yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII,2018).

Pola konsumsi masyarakat yang semakin banyak menggunakan platform digital untuk memenuhi kebutuhan konsumsi mereka turut mempengaruhi industry travel. Sector online travel ini menyumbang lebih dari 38 persen dari perkiraan pertumbuhan ekonomi internet di Asia Tenggara. Sector penjualan online untuk travel memang semakin diminati masyarakat Indonesia. Industri perdagangan digital, terutama di sector online travel diprediksi akan terus bertumbuh. Pada tahun 2015 mencapai 34% dari total nilai transaksi layanan perjalanan dan meningkat menjadi 41% pada 2018, lalu diprediksi kembali tumbuh mencapai 57% pada tahun 2025 (Google dan Temasek, 2108). Lembaga Riset Alvara Research Center yang merupakan lembaga penelitian asal indonesia yang mengedepankan akurasi, dalam hasil penelitian berjudul Perilaku dan Preferensi Konsumen Milenial Indonesia terhadap Aplikasi E-

Commerce 2019. Survei tersebut dengan tujuan mengidentifikasi ciri khas dari kaum milenial, serta bagaimana ciri-ciri tersebut dapat terlihat dari preferensi e-commerce pada saat transaksi e-commerce terus meningkat dan cepat. Survei dilakukan terhadap 1.204 responden dari generasi milenial mencatat merek dagang Traveloka menjadi platform pemesanan tiket dan hotel paling diminati milenial di rentang usia 17 hingga 39 tahun. Traveloka muncul sebagai aplikasi yang paling sering digunakan oleh 79 persen responden, sebuah angka mendominasi aplikasi pemesanan hotel dan tiket lainnya seperti Tiket.com (8,9%), Blibli.com (5,6%), KAI Acces (3,2%) dan Airy (2,4%) .

Traveloka adalah *online travel agent* di Indonesia yang menyediakan layanan pemesanan tiket pesawat, hotel, tiket kereta, paket pesawat dan hotel, aktivitas dan rekreasi, produk-produk konektivitas, transportasi bandara dan bus. Aplikasi mobile Traveloka telah disediakan di situs *Google Play* dan diunduh lebih dari 30 juta kali, menjadikannya aplikasi mobile paling populer di Asia Tenggara. Hal yang menarik dari aplikasi *google play* ialah adanya sebuah fitur yang dapat memberikan penilaian dan ulasan dari para penggunanya.

Dari banyaknya ulasan yang diberikan oleh para penggunanya dapat digunakan sebagai suatu tolak ukur yang efektif dan efisien untuk menemukan informasi tentang emosi yang terkandung dalam ulasan terhadap aplikasi Traveloka. Ulasan sendiri berupa teks yang digunakan untuk berkomunikasi dan menyampaikan informasi. Tidak hanya memuat informasi, teks juga dapat mengekspresikan emosi (Hirat & Mittal, 2015; Mohammad & Kiritchenko, 2015).

Text mining adalah teknologi yang mampu menganalisis data teks semi terstruktur maupun tidak terstruktur, sedangkan data mining mengolah data yang terstruktur (Han & Kamber, 2012). Dengan cara yang dianalogikan dengan *data mining*, *text mining* berupaya mengekstraksi informasi yang berguna dari sumber data melalui identifikasi dan eksplorasi pola yang menarik.

Ada beragam teknik klasifikasi dokumen, di antaranya adalah Naïve Bayes *classifier*, *Decision Trees*, dan *Support Vector Machines*. paling populer digunakan dalam pengklasifikasian dokumen sekarang ini adalah metode Naïve Bayes Classifier (Natalius, 2011). Metode Naïve Bayes Classifier mempunyai kecepatan dan akurasi yang tinggi ketika diaplikasikan dalam basis data yang besar dan data yang beragam (Larose, 2006). Hal serupa juga diungkapkan oleh (A Syarifah, 2015) dalam penelitiannya, yaitu metode Naïve Bayes Classifier memiliki beberapa kelebihan antara lain, sederhana, cepat dan berakurasi tinggi.

Berdasarkan dari penelitian yang di lakukan oleh Dinar Ajeng Kristiyanti, Akhmad Hairul Umam, Mochamad Wahyudi, Mochamad Wahyudi dan Linda Marlinda yang berjudul “Comparison of SVM & Naïve Bayes Algorithm for Sentiment Analysis Toward West Java Governor Candidate Period 2018-2023 Based on Public Opinion on Twitter”, membuktikan bahwa Naive Bayes memiliki presentase akurasi terbesar dibanding Support Vektor Machine. Penelitian yang bertujuan untuk membandingkan algoritma terbaik pada sentiment analisis ini menggunakan dua jenis kelas sentimen yaitu positif dan negatif. Kesimpulan dari penelitian ini mengatakan bahwa Akurasi tertinggi Naïve adalah 94% sedangkan

Dukungan Algoritma Vector Machine hanya menghasilkan akurasi tertinggi dari 75,50%. Perbedaan akurasi cukup jauh, yaitu 18,50%.

Menurut Daniel Goleman (2002 : 411) emosi merujuk pada suatu perasaan dan pikiran yang khas, suatu keadaan biologis dan psikologis dan serangkaian kecenderungan untuk bertindak. Emosi pada dasarnya adalah dorongan untuk bertindak. Biasanya emosi merupakan reaksi terhadap rangsangan dari luar dan dalam diri individu.

Pada tahun 2014, Rachael E. Jack, Oliver G. B. Garrord, dan Phillipe G. Schyns dari Institute of Neuroscience and Psychology, University of Glasglow melakukan penelitian tentang emosi dasar menggunakan 3 metode, yaitu *perceptual expectation modeling*, *information theory*, *Bayesian Classifier*. Hasil dari penelitian tersebut adalah 4 emosi dasar, yaitu senang, sedih, terkejut, marah. Dengan menggunakan metode TF-IDF (*Term Frequency – Inverse Document Frequency*) dan Naïve Bayes bisa menentukan emosi apa yang ada pada ulasan pada aplikasi E – Ticketing Traveloka. Hasil dari penelitian ini diharapkan dapat mempermudah, membantu serta memberikan informasi tentang emosi yang terkandung dalam ulasan tentang aplikasi *E – Ticketing* Traveloka yang ada pada situs Google Play.

1.2. Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, terdapat permasalahan yang akan diangkat dalam skripsi ini yaitu bagaimana membangun model klasifikasi terbaik yang dapat menganalisis emosi pada ulasan aplikasi Traveloka yang di ambil dari situs *google play*.

1.3. Batasan Masalah

Batasan dan ruang lingkup permasalahan dalam skripsi ini yaitu:

- a. Data *ulasan* aplikasi Traveloka yang diambil dari *google play* memiliki rentang waktu bulan November 2019 hingga bulan Januari 2020
- b. Data *ulasan* yang diambil hanya data *ulasan* yang berbahasa Indonesia.
- c. Emosi yang diklasifikasikan berdasarkan hasil penelitian yang dilakukan Rachael E. Jack, Oliver G. B. Garrord, dan Phillipe G. Schyns (2014) yaitu senang, sedih, terkejut, dan marah.
- d. Tidak menganalisis kata atau ulasan sarkasme

1.4. Tujuan

Sesuai dengan rumusan masalah yang diangkat, tujuan dari penulisan skripsi ini bertujuan yaitu untuk membangun model pengklasifikasian terbaik untuk dapat menganalisis emosi pada data ulasan pengguna aplikasi Traveloka dengan menggunakan metode *Naïve bayes*.

1.5. Manfaat

Adapun manfaat yang ingin dicapai dan diperoleh dari penulisan skripsi ini adalah Dapat menerapkan analisis emosi pada data ulasan pengguna aplikasi Traveloka dengan menggunakan metode *Naïve bayes* dan mengetahui minat masyarakat terhadap aplikasi Traveloka dengan melihat sentimen emosi yang dihasilkan.

1.6. Sistematika Penulisan

Dalam laporan skripsi ini, pembahasan disajikan dalam lima bab dengan sistematika pembahasan sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisikan tentang latar belakang, perumusan masalah, batasan masalah, tujuan, manfaat, metode penelitian dan sistematika penulisan yang digunakan dalam laporan skripsi ini.

BAB II TINJAUAN PUSTAKA

Bab ini membahas tentang penelitian sebelumnya, teori teori mengenai konsep *Explatory Analysis Data (EDA)*, emosi, *text mining*, analisis sentiment, *Term Frequency – Inverse Document Frequency*, *Naïve Bayes*, *Evaluasi model*, Google Play, E - Ticketing, Traveloka dan python.

BAB III METODOLOGI PENELITIAN

Bab ini berisi dijelaskan tentang kebutuhan data penelitian, perancangan model dan perancangan sistem.

BAB IV : HASIL DAN PEMBAHASAN

Pada bab ini dijelaskan tentang hasil pengujian model dan implementasi sistem.

BAB V : KESIMPULAN DAN SARAN

Pada bab ini dijelaskan tentang kesimpulan dari penelitian dan saran untuk penelitian selanjutnya.

DAFTAR PUSTAKA

Pada bagian ini akan dipaparkan tentang sumber-sumber literatur yang digunakan dalam pembuatan laporan ini. Berisi tentang penjelasan dan pembahasan pendukung skripsi guna memahami isi laporan.