

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pesatnya perkembangan teknologi informasi dan komunikasi menyebabkan bahwa negara-negara maju dan berkembang telah mengubah strategi dalam ekonomi dan bisnis. Teknologi menjadi alat yang membantu kebutuhan manusia, dengan teknologi, apapun dapat dilakukan dengan lebih mudah. Begitu pentingnya peran teknologi inilah yang mulai membawa peradaban memasuki ke era digital. Banyak perusahaan dengan tergesa-gesa menggunakan teknologi informasi dan komunikasi termasuk perusahaan yang bergerak dibidang perdagangan (Karami, 2014). Pemanfaatan teknologi informasi dan komunikasi di bidang perdagangan berupa penjualan dan pembelian barang secara online. Sebuah survei yang diselenggarakan Asosiasi Penyelenggaraan Jasa Internet Indonesia (APJII) menunjukkan bahwa jumlah pengguna internet di Indonesia meningkat tiap tahunnya hingga akhir tahun 2018 mencapai 171,17 juta orang, dari total penduduk sebanyak 264,16 juta orang. Perkembangan tersebut berdampak pada berbagai bidang. Salah satunya ialah maraknya kegiatan berbelanja melalui media internet. Menurut survei yang dilakukan oleh APJII 47,6% dari pengguna internet di Indonesia mengetahui internet sebagai tempat jual beli barang dan jasa, dan 41,2% pernah melakukan transaksi secara online. Istilah yang digunakan dalam transaksi jual beli secara online adalah e-commerce (APJII, 2018).

Indonesia telah menjadi pasar terbesar e-commerce di Asia Tenggara, penjualan online Indonesia mencapai US\$ 1,1 miliar, lebih tinggi dari Thailand

dan Singapura. Namun, jika dibandingkan dengan total perdagangan retail, penjualan e-commerce di Indonesia hanya menyumbang 0,07 persen. Artinya, pasar e-commerce Indonesia berpeluang untuk tumbuh semakin besar. Apalagi dengan jumlah penduduk dan tingkat produk domestik bruto (PDB) terbesar di ASEAN. Euromonitor memperkirakan rata-rata pertumbuhan tahunan (CAGR) penjualan online Indonesia selama 2014-2017 sebesar 38% (Iprice, 2019). APJII pada 2018 juga melakukan survei dua konten internet komersial teratas yang sering digunakan untuk membeli barang atau jasa online yakni Shopee 11,2% dan Bukalapak 8,4%. Bukalapak dan Shopee merupakan model e-commerce Customer-To Customer (C2C). Model ini yang paling banyak diterapkan dan di temukan pada situs e-commerce di Indonesia. Pasar Customer-To Customer (C2C) saat ini masih dominan di pasar ritel online Indonesia. Bukalapak dan Shopee sudah memiliki aplikasi mobile phone dan website, yang memberikan akses kepada pengguna untuk memberikan ulasan terkait dua e-commerce tersebut (Iprice, 2019).

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

Pelanggan yang merasa tidak puas dengan layanan atau produk yang ditawarkan oleh sebuah perusahaan biasanya akan menuliskan keluhannya di media sosial. Di sisi lain, ada juga pelanggan yang merasa puas, yang mengekspresikan sikap positif mereka terhadap sebuah produk di media sosial. Tanpa disadari, opini-opini pelanggan yang dituliskan di media sosial akan memberikan pengaruh pada calon pelanggan. Akan tetapi, memantau opini dari masyarakat juga bukanlah hal yang mudah. Opini yang dimuat di media sosial jumlahnya terlalu banyak untuk diproses secara manual. Oleh sebab itulah, diperlukan sebuah metode atau teknik khusus yang mampu mengkategorikan review-review tersebut secara otomatis, apakah termasuk positif atau negatif (Potdar, Patil, Bagla, Pandey, & Jadhav, 2016).

Sentiment Analysis (SA) adalah cara mengidentifikasi dan mengelompokkan polaritas teks yang diberikan ditingkat dokumen, kalimat dan frasa (Kang, Yoo, & Han, 2012). Teknik ini digunakan di banyak bidang seperti e-commerce, perawatan kesehatan, hiburan dan politik, untuk beberapa nama. Sebagai contoh, Analisis Sentimen berguna bagi perusahaan untuk memantau opini konsumen mengenai produk mereka, dan bagi konsumen untuk memilih produk terbaik berdasarkan opini publik. Tantangan utama analisis sentimen adalah ulasan umumnya ditulis dalam bahasa informal, pendek pesan menunjukkan isyarat terbatas tentang sentimen serta akronim dan singkatan banyak digunakan. Secara umum, teknik klasifikasi dasar seperti Naïve Bayes (Multinomial and Bernoulli), Logistic Regression, SGD (*Stochastic Gradient Descent*), Linear SVM (Support Vector Machine) dan Random Forest (Mohaiminul & Sultana, 2018)

Penelitian yang dilakukan oleh Saraswati (2011) mengklasifikasikan opini dengan menggunakan metode NBC dan SVM. Dari hasil penelitiannya, Saraswati menyatakan bahwa metode SVM memiliki tingkat akurasi yang lebih tinggi daripada metode NBC untuk pengujian data opini positif, sedangkan metode NBC menunjukkan hasil yang lebih baik jika digunakan dalam pengujian data opini negatif.

Putranti dan Winarko (2014) tentang analisis sentimen dalam penelitian ini proses klasifikasi dokumen tekstual dibagi kedalam dua kelas, yaitu kelas sentimen positif dan negatif. Data opini diperoleh dari jejaring sosial Twitter berdasarkan query dalam Bahasa Indonesia. Penelitian ini bertujuan untuk menentukan sentimen publik terhadap objek tertentu yang disampaikan di Twitter dalam bahasa Indonesia, sehingga membantu usaha untuk melakukan riset pasar atas opini publik. Algoritma Maxent digunakan untuk POS tagger dan SVM digunakan untuk membangun model klasifikasi. Implementasi klasifikasi diperoleh akurasi 86,81 % pada pengujian 7 fold cross validation untuk tipe kernel Sigmoid. Pelabelan kelas secara manual dengan POS tagger menghasilkan akurasi 81,67%.

Penelitian yang dilakukan oleh Alamsyah dkk (2017) yang berkaitan dengan analisis sentimen untuk mengetahui bagaimana perbandingan kepuasan pelanggan terhadap tiga situs e-commerce yang sering dikunjungi di Indonesia yaitu Bukalapak, Tokopedia dan Elevenia. Data yang didapatkan dari media sosial twitter, dalam penelitiannya Naïve Bayes Classifier digunakan sebagai teknik klasifikasi dengan pembobotan TF-IDF, sedangkan untuk melakukan validasi dan

evaluasi pada klasifikasi teks Naïve Bayes dilakukan menggunakan K-fold cross validation dan confusion matrix. Hasil penelitian menunjukkan bahwa sentimen negatif terhadap ketiga situs e-commerce tersebut lebih mendominasi di media sosial, dan situs e-commerce dengan sentimen negatif paling tinggi ialah Bukalapak kemudian Tokopedia dan Elevenia (Alamsyah & Saviera, 2017).

Mohaiminul Islam dan Sultana tahun 2018 juga melakukan komparasi beberapa metode untuk klasifikasi analisis sentimen. Metode yang digunakan antara lain Multinomial Naïve Bayes, Bernouli Naïve Bayes, Logistic Regression, Linier SVM, *Stochastic Gradient Descent* (SGD) dan Random Forest. Data yang digunakan adalah dataset IMBD dan Amazon, dengan atribut rating dan review. Berdasarkan hasil yang didapat untuk kedua dataset Linier SVM yang memiliki tingkat akurasi paling tinggi, yang selanjutnya *Stochastic Gradient Descent* memiliki tingkat akurasi tertinggi kedua (Mohaiminul & Sultana, 2018).

Gusriani, dkk (2016) melakukan analisis sentimen berdasarkan komentar publik terhadap toko online seperti Zalora dan Berry Benka pada media sosial facebook. Metode klasifikasi Naive Bayes, K-NN dan Decision Tree digunakan untuk membandingkan hasil prediksi klasifikasi yang terbaik. Hasil analisis pengujian menunjukkan Naive Bayes, memiliki kestabilan akurasi setelah diuji dengan beberapa nilai Frequent Itemset. Naive Bayes memiliki rata-rata akurasi 90,3%.

Moh dkk pada tahun 2015 juga melakukan penelitian mengenai multi-tier classification untuk analisis sentimen dengan menggunakan 150.000 review film yang menggunakan beberapa metode antara lain Naïve Bayes, SVM (Support

Vector Machine), Random Forest, dan *Stochastic Gradient Descent*. Berdasarkan hasil yang didapat *Stochastic Gradient Descent* memiliki akurasi paling tinggi saat dilakukan multi-tier sentimen analisis (Moh, Gajjala, Gangireddy, & Moh, 2016).

Stochastic Gradient Descent merupakan salah satu tipe dari Gradient Descent. Penerapan SGD dalam proses klasifikasi memiliki kelebihan mampu mengoptimasi dalam data yang begitu besar sebab *Stochastic Gradient Descent* menemukan gradien fungsi biaya dari satu contoh pada setiap iterasi, bukan jumlah gradien fungsi biaya dari semua contoh, selain itu SGD mampu meningkatkan kinerja generalisasi masalah berskala besar, Mampu diterapkan pada data dengan jumlah besar dan memiliki missing value dan waktu komputasi relatif lebih singkat meskipun datanya besar dan dari beberapa penelitian memiliki tingkat akurasi yang lebih tinggi dari algoritma klasifikasi pada analisis sentimen lainnya (Sharma, 2018). Oleh karena itu pada penelitian ini penulis akan melakukan Analisis Sentimen Terhadap Review E-Commerce dengan Metode *Stochastic Gradient Descent* dengan studi kasus Shopee dan Bukalapak.

1.2.Perumusan Masalah

1. Bagaimana gambaran umum data review tentang Shopee dan Bukalapak berdasarkan situs Google Play Store?
2. Bagaimana hasil mengklasifikasikan teks berbahasa Indonesia mengenai review pengguna Shopee dan Bukalapak berdasarkan situs Google Play Store menggunakan metode *Stochastic Gradient Descent*?

3. Bagaimana hasil dari ulasan negatif yang di sampaikan oleh konsumen terhadap penyedia e-commerce?

1.3.Tujuan Penelitian

1. Mengetahui gambaran umum tentang review Shopee dan Bukalapak berdasarkan situs Google Play Store
2. Mengetahui hasil klasifikasi teks berbahasa Indonesia mengenai review pengguna Shopee dan Bukalapak berdasarkan situs Google Play Store menggunakan metode *Stochastic Gradient Descent*.
3. Mengetahui informasi yang di sampaikan oleh konsumen terhadap penyedia e-commerce berdasarkan hasil dari ulasan negatif

1.4.Manfaat Penelitian

1. Bagi Pembaca
Dapat menambah dan memperkaya pengetahuan tentang implementasi statistika dalam dunia perdagangan online serta sebagai sarana informasi bagi pembaca dan pelengkap referensi bagi pihak-pihak yang memerlukan.

2. Bagi Universitas Muhammadiyah Semarang
Dapat digunakan untuk bahan referensi dalam pengaplikasian ilmu statistika khusus untuk pengolahan data teks dengan menggunakan metode *text mining* kedalam dunia perdagangan.

3. Bagi penyedia layanan e-commerce

Manfaat Pengklasifikasian data komentar Shopee dan Bukalapak Bukalapak dapat memudahkan pihak perusahaan dalam mengetahui persepsi pengguna jasa dalam bentuk opini negatif dan opini positif, sehingga dapat dijadikan sebagai acuan dalam upaya menjaga kualitas dan memperbaiki kekurangan serta evaluasi ke arah yang lebih baik.

1.5.Batasan Penelitian

Penelitian yang dilakukan hanya untuk mengklasifikasikan data hasil pelabelan sentimen dan di evaluasi dengan akurasi. Selain itu waktu yang digunakan adalah data bulan Juni 2019 berasal dari data Google Play Store.

