

BAB I

PENDAHULUAN

I.1. Latar Belakang

Teknologi adalah suatu sarana atau sistem yang berfungsi untuk memberikan kenyamanan serta kemudahan bagi manusia. Pada zaman ini perkembangan teknologi semakin mengalami perkembangan yang sangat pesat. Hal ini disebabkan oleh keanekaragaman kebutuhan masyarakat yang sangat didukung oleh Jenis-jenis teknologi dan juga semakin ketatnya persaingan penyedia teknologi. Dengan berkembangnya teknologi Informasi sangat memudahkan masyarakat dalam memenuhi kebutuhan sehari-hari. Contoh teknologi yang nyata di era sekarang sangat banyak, mulai dari teknologi yang kita sering gunakan seperti komputer, laptop, handphone, mesin-mesin, mobil, motor, dan lainnya. Seiring perkembangan teknologi yang semakin maju tentu akan lebih banyak teknologi yang muncul. (Muhson, 2010)

Perkembangan teknologi telah membuat pergeseran perilaku pelanggan dari pembelian melalui offline shop menjadi pembelian melalui online shop atau melalui e-commerce. Perubahan tersebut memengaruhi keputusan pembelian konsumen dalam membeli suatu produk. Hal tersebut banyak dipengaruhi oleh persepsi mereka terhadap harga, produk, promosi, tempat (bauran pemasaran) yang telah diterapkan oleh perusahaan selama ini. (Kotler *et al.*, 2016) E-commerce adalah tentang menjual online atau kemampuan untuk bertransaksi secara online, termasuk ritel, perbankan online dan belanja – yang melibatkan transaksi dimana pembeli

benar-benar membeli dan berbelanja. (Chaffey *et al.*, 2019) Munculnya E-Commerce sangat bermanfaat bagi keberlangsungan hidup masyarakat. Contoh manfaat bagi konsumen adalah Konsumen dapat berbelanja dengan lebih mudah selama 24 jam sehari tanpa harus datang langsung ke Toko, konsumen dapat melihat berbagai produk dengan harga yang paling tepat. Konsumen dapat membeli produk dan jasa dengan biaya yang lebih mudah setelah melakukan perbandingan dengan berbagai e-commerce. (Sarifudin & Amelia, 2020)

Perubahan perilaku masyarakat dalam E-commerce merupakan tantangan tersendiri bagi perusahaan agar dapat memenuhi kebutuhan dan keinginan konsumennya. Dengan bertambahnya jenis-jenis E-Commerce tentunya sangat mempengaruhi daya saing antar E-Commerce. Hal ini menyebabkan setiap pebisnis yang bergabung dengan E-Commerce mempunyai cara tersendiri untuk meningkatkan daya tarik konsumen pada E-Commerce tersebut. Hal-hal yang selalu diperhatikan pebisnis E-Commerce untuk meningkatkan daya tarik konsumen yaitu memperhatikan kualitas pelayanan terhadap Konsumen, menggunakan News letter, selalu mengerjakan upsell produk, mengerjakan retargeting iklan. Namun walaupun demikian tidak jarang konsumen meragukan produk dari E-Commerce dikarenakan konsumen tidak dapat mencoba atau melihat secara langsung produk atau jasa dari Produsen E-Commerce yang bersangkutan. (Supradono & Hanum, 2011) Berdasarkan hasil analisis yang telah dilakukan sebagian besar setiap konsumen yang pernah berbelanja di E-Commerce akan memberikan review terhadap produk yang telah dibeli. Review maupun rating juga sangat penting untuk meningkatkan frekuensi pelanggan, karena review dan rating

dari pelanggan dapat memberikan penilaian yang lebih bersifat akurat dan juga emosional karena diberikan oleh sesama pelanggan sehingga memiliki nilai kepercayaan yang lebih tinggi. Oleh sebab itulah rating sebuah toko online membutuhkan perhatian khusus dari pengelola toko online untuk meningkatkan pelanggannya. (Farki, 2016)

Rating adalah bagian dari review yang menggunakan bentuk simbol bintang daripada bentuk teks dalam mengekspresikan pendapat dari pelanggan. Rating dapat diartikan sebagai penilaian dari pengguna pada preferensi suatu produk terhadap pengalaman mereka mengacu pada keadaan psikologis dan emosional yang mereka jalani saat berinteraksi dengan produk virtual dalam lingkungan dimediasi. (Kim & Park, 2017) Namun kepercayaan pelanggan terhadap rating dapat terhambat karena inkonsistensi jawaban seperti review yang memiliki teks negatif dengan nilai rating Positif. Adanya jawaban yang berbeda seperti rating bagus namun review buruk dan sebaliknya dapat membuat orang lain bingung. Untuk itu dibutuhkan penelitian untuk menemukan suatu cara yang akurat untuk mensinkronisasi antara review konsumen dengan rating yang diberikan karena pada dasarnya rating harus sesuai dengan review untuk memudahkan konsumen dalam memilih produk dari E-Commerce. (Amelia *et al.*, 2021)

Topik dalam penelitian ini adalah bagaimana mengidentifikasi teks atau komentar-komentar pada online shop yang akan di analisis dan divisualisasikan sehingga dapat dijadikan sebagai nilai rating dari produk yang dibelanjakan tersebut. Proses mengidentifikasi teks pada sebuah database disebut dengan istilah *text mining*. Text mining adalah proses mengambil Informasi berkualitas tinggi dari

teks informasi berkualitas tinggi. *Text mining* bertujuan untuk mendapatkan informasi yang berguna dari sebuah dokumen. (Ipmawati, 2016)

Pada fungsi *data mining*, dibutuhkan algoritma beserta metode yang sesuai dalam pengolahan input menjadi output yang diinginkan. Atas dasar inilah, terdapat banyak algoritma *data mining* untuk setiap fungsi. Misalnya, algoritma fungsi klasifikasi terdiri dari *Statistical-Based Algorithms*, *Distance-Based Algorithms*, *Decision Tree-Based Algorithms*, *Neural Network-Based Algorithms*, dan *Rule-Based Algorithms*. (Rasheed et al., 2020)

(Gunawan et al., 2018) melakukan penelitian dengan menggunakan metode algoritma Naive Bayes untuk Analisis sentimen pada Ulasan produk. Pengukuran kinerja didasarkan atas jumlah data dan jumlah atribut. Hasilnya, diperoleh bahwa sistem analisis yang dibentuk belum sepenuhnya relevan dalam memprediksi kelas sentimen yang sesuai terhadap pemberian kelas dari supervisor. Dan jumlah data latih dalam sistem analisis sentimen memiliki pengaruh terhadap prediksi program.

Penelitian analisis sentimen juga dilakukan oleh Andi Nurul Hidayat yaitu Analisis sentimen menggunakan algoritma Support Vector Machine dan Naive Bayes. Dari hasil analisis dapat disimpulkan bahwa penerapan metode algoritma SVM menunjukkan akurasi yang lebih besar dibandingkan dengan Naive Bayes. Pada masing masing data yang telah di analisis standar dengan teknik pengukuran Precision, Recall. (Hidayat, 2015)

Berdasarkan latar belakang tersebut, penelitian ini mencoba melakukan analisis pada komentar-komentar hasil review pada belanja *Online Shopping* dengan menggunakan dua metode algoritma untuk mendapatkan hasil prediksi yang

lebih baik, yaitu algoritma *Naïve Bayes*, dan *Support Vector Machine* (SVM) yang merupakan bagian dari algoritma *data mining* yang terbaik. Dan melakukan pengembangan dengan Visualisasi sentimen untuk menciptakan kerapian data berdasarkan komentar yang paling sering muncul. Adapun sumber data set adalah dari situs Kaggle dengan keyword pencarian Lazada. Pengukuran akurasi menghasilkan *confusion matrix* untuk penilaian *precision*, *recall*. Atas dasar inilah penelitian ini mengambil judul: “Analisis Pendekatan *Machine Learning* Untuk Memprediksi Rating belanja Online Pada Aplikasi Online Shopping Menggunakan Algoritma *Naïve Bayes* Dan *Support Vector Machine*”

I.2. Rumusan Masalah

Pada penelitian ini, rumusan masalah meliputi:

1. Bagaimana menangani dataset yang tidak konsisten?
2. Bagaimana menyesuaikan data komentar dengan rating pada *online shopping*?
3. Bagaimana menganalisa *performance* dari algoritma *Naïve Bayes* dan SVM dari sudut pandang berdasarkan *confusion matrix* pada permasalahan *text mining*?

I.3. Tujuan

Adapun tujuan penelitian ini yaitu:

1. Memperoleh hasil klasifikasi dari data penelitian, yaitu berupa data komentar yang dianalisis secara otomatis dan dikelompokkan menjadi rating menggunakan algoritma *Naïve Bayes* dan SVM.

2. Menganalisis hasil akurasi dari algoritma klasifikasi *Naïve Bayes*, dan SVM berdasarkan *confusion matrix*.
3. Memperoleh data visualisasi berdasarkan data komentar.

I.4. Manfaat

Manfaat yang diharapkan dari hasil penelitian ini adalah sebagai berikut:

1. Dapat mengetahui lebih dalam tentang tahapan klasifikasi pada masalah *text mining*.
2. Dapat mengetahui penilaian akurasi algoritma *Naïve Bayes* dan SVM berdasarkan *confusion matrix*, meliputi *precision*, *recall*, *F-measure*, dan *success rate*.
3. Hasil penelitian dapat menjadi rujukan para pembaca dalam memahami algoritma *Naïve Bayes*, SVM, terutama dalam bagian akurasi.
4. Hasil Analisis dan Klasifikasi Algoritma SVM dan *Naïve Bayes* dapat membantu Konsumen untuk melihat hasil rating yang di analisis berdasarkan data komentar pada Platform Lazada sehingga proses belanja lebih efisien.

I.5. Batasan Masalah

Dalam menganalisis komentar-komentar menggunakan algoritma *Naïve Bayes*, dan SVM dan juga visualisasi dengan menggunakan Bahasa program *Phyton* dilakukan beberapa batasan sebagai berikut:

1. Data set yang digunakan adalah komentar yang berupa opini maupun review belanja online platform Lazada yang diambil dari situs Kaggle.

2. Data set yang digunakan berupa teks komentar-komentar pada online shop Lazada dalam menganalisis *review* maupun komentar dan mengelompokkannya menjadi rating menggunakan algoritma *Naïve Bayes*, dan SVM.

I.6. Sistematika Penulisan

Sistematika proposal tesis yang akan dilakukan dalam penelitian ini terdiri dari 3 bab, yaitu:

BAB I PENDAHULUAN

Pendahuluan akan membahas tentang latar belakang dari masalah yang akan di angkat beserta rumusan masalah, tujuan penelitian, dan batasan masalah serta sistematika penulisan yang akan diuraikan dalam penyusunan laporan penelitian ini.

BAB II TINJAUAN PUSTAKA

Tinjauan pustaka menjelaskan tentang teori-teori yang akan diuraikan mengenai komentar, kritikan, saran pada aplikasi Online Shop, beserta perangkat analisa sistem beserta metode analisa yang digunakan.

BAB III METODOLOGI PENELITIAN

Metodologi penelitian menjelaskan rangkaian tahapan dalam pembuatan sistem yang terdiri dari pengumpulan data, analisis data,

perancangan sampai implementasi beserta analisa yang akan dilakukan.

BAB IV HASIL

Hasil penelitian yaitu berupa pemaparan hasil pelaksanaan metode/teknik penelitian dan menyajikan data yang mendukung hasil tersebut. Penyajian data dan penjelasannya dilakukan secara berurutan dan lengkap menggunakan teks dan ilustrasi lainnya (misalnya, tabel dan gambar) sebagai pendukung hasil.

BAB V PEMBAHASAN

Pembahasan penelitian yaitu menunjukkan makna dari hasil yang diperoleh untuk menjawab pertanyaan atau masalah penelitian. Fungsi lainnya adalah untuk menerangkan pemahaman baru yang diperoleh dari hasil penelitian, yang diharapkan berguna dalam pengembangan keilmuan berikutnya. Pada penelitian tingkat lanjut, hasil pembahasan diharapkan memiliki kontribusi penulis terhadap pengembangan keilmuan kedepannya. Dalam penelitian ini, hasil pembahasan bukan pengembangan keilmuan secara substansial, tetapi cukup terhadap pemahaman secara personal dalam implementasi konsep atau teori.

BAB VI KESIMPULAN DAN SARAN

Kesimpulan dan saran berisi hasil penelitian klasifikasi terhadap komentar maupun review yang dikelompokkan menjadi rating menggunakan metode *Naive Bayes* dan SVM, Visualisasi dan simulasi interaktif dengan menggunakan Phyton. Bagian kesimpulan berisi jawaban atas rumusan masalah pada penelitian dan pada bagian saran berupa pengembangan penelitian dan kendala yang dihadapi dalam menyusun penelitian sehingga dapat memberikan masukan bagi penelitian selanjutnya.