

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Jaringan media sosial pada saat ini terus berkembang dan berdampak pada industri perhotelan. Pelanggan dan *traveler* telah memposting hasil *review* secara online untuk menunjukkan tingkat kepuasan mereka terhadap hotel dan berbagi pengalaman terkait hotel yang dikunjungi dengan pelanggan lain yang ada di seluruh belahan dunia. Situs web yang bergerak dalam pariwisata dan perhotelan berkembang pesat secara online seperti *trip advisor* [1]. *Trip advisor* merupakan *platform* penyedia layanan perjalanan dan pemesanan hotel. Data yang diperoleh dalam peneliti ini berisikan dataset hotel yang ada di kota *Mussoorie*, data yang digunakan berasal dari *platform* perjalanan online asal amerika yaitu: *trip advisor* yang ada di *kaggle*. Dalam hal ini kami menggunakan teknik analisis sentimen untuk mengkategorikan opini pengguna yang bernilai negatif maupun positif dengan bantuan kecerdasan buatan yaitu *machine learning*.

Banyak sekali informasi yang ditampilkan di website *trip advisor*, salah satunya adalah klasifikasi hotel berdasarkan *review* pelanggan. Ada beberapa klasifikasi hotel yang memiliki rating tertinggi hingga terendah. *Review* hotel sangat perlu untuk para *traveler*, dengan adanya *review* tersebut maka terciptanya sebuah rating sehingga para *traveler* dapat memutuskan hotel mana yang sesuai dengan keinginan mereka untuk menginap. Permasalahan yang dapat ditimbulkan ketika pelanggan memberikan hasil *review* mereka terhadap hotel yang pernah dikunjungi bisa berisi komentar yang positif atau negatif, hal tersebut dapat menyulitkan dan membuat bingung para pengunjung yang ingin memesan hotel melalui *platform trip advisor*. Oleh karena itu perlu dilakukan analisis sentimen terhadap hasil *review* para pelanggan agar pengunjung tidak sulit dalam menentukan hotel mana yang bagus atau tidak sesuai *review* oleh pengunjung sebelumnya.

Telah banyak penelitian-penelitian sebelumnya yang sudah melakukan analisis sentimen *review* hotel, sebagai contoh “*Sentiment analysis on hotel reviews using Multinomial Naïve Bayes classifier*” dan “*Sentiment Analysis on Chinese Hotel Reviews with Doc2Vec and Classifiers*” [2][3]. Penelitian yang dilakukan menggunakan data hotel yang ada di *platform trip advisor* untuk menganalisis sentimen sebuah *review* hotel yang dilakukan oleh pelanggan dengan menggunakan

metode yang berbeda. Untuk penelitian pertama menggunakan metode multinomial naïve bayes classifier dengan hasil preprocessing+BoW (bag of words) sebesar 90.927%, preprocessing+feature selection-1 sebesar 91,414% dan preprocessing+featur selection-2 sebesar 91.233%.

Sedangkan penelitian kedua menggunakan metode Doc2vec dan pengklasifikasian algoritma dan mendapatkan hasil akhir algoritma svm (*Support Vector Machine*) sebesar 79.5%, *algoritma logistic regression* sebesar 78.83% dan naïve bayes sebesar 73.42%.

Hasil perbandingan penelitian pertama dan kedua dapat dilihat pada tabel dibawah ini:

Tabel 1 hasil perbandingan penelitan pertama dan kedua

Penelitian I	Hasil penelitian	Penelitian II	Hasil penelitian
preprocessing+BoW	90.927%	Algoritma Svm	79.5%
preprocessing+feature selection -1	91,414%	Algoritma Logistic Reegression	78.83%
preprocessing+featur selection-2	91.233%	Naïve Bayes	73.42%

Pada perbandingan penelitian sebelumnya didapatkan hasil akurasi tertinggi dalam review hotel sebesar 91,414% oleh karena itu berlu dilakukan pengembangan dalam melakukan analisis review hotel dengan menggunakan algoritma yang berbeda agar mendapatkan hasil akurasi lebih baik dari penelitian sebelumnya.

Berdasarkan latar belakang yang telah dipaparkan, penulis tertarik untuk melakukan peneltia dengan judul “**Analisis Perbandingan *Decision Tree*, *Support Vector Machine*, dan *Xgboost* dalam Mengklasifikasi *Review* Hotel Pada *Trip Advisor***”.

1.2 Rumusan masalah

Rumusan masalah yang dapat dibuat berdasarkan latar belakang yang telah diuraikan adalah sebagai berikut:

1. Bagaimana pelanggan memberikan hasil *review* hotel?
2. Apakah hasil *review* oleh pelanggan bernilai positif atau negatif?
3. Bagaimana *platform trip advisor* melakukan rating?

4. Apakah algoritma pengolahan teks dapat menganalisis hasil review hotel secara otomatis?

1.3 Batasan masalah

Agar penelitian yang dilakukan menjadi lebih spesifik sesuai perumusan masalah tersebut maka batasan yang ditentukan adalah sebagai berikut:

1. Dataset yang diambil dan dianalisis *Trip advisor* dataset dari *kaggle*.
2. Metode yang digunakan adalah metode pengolahan teks.
3. Algoritma yang dipakai adalah algoritma sentimen analisis.

1.4 Tujuan penelitian

Tujuan diadakannya penelitian yaitu untuk:

1. Mengetahui apakah metode analisis sentimen mampu secara otomatis menganalisis hasil review pelanggan hotel.
2. Mengetahui akurasi yang diperoleh dari algoritma analisis sentimen untuk melakukan klasifikasi hasil review pelanggan yang bernilai positif atau negatif.
3. Menentukan review mana yang bernilai positif dan negatif berdasarkan penilaian yang objektif menggunakan metode sentimen analisis.

1.5 Manfaat penelitian

Manfaat yang dapat diberikan penelitian ini adalah:

1. Membantu menganalisis algoritma sentimen mengatasi permasalahan pada hasil *review* pelanggan hotel di *platform trip advisor*.
2. Menjadi referensi bagi peneliti mendatang bila akan melakukan penelitian dengan tema terkait.
3. Membantu pemilik hotel dalam meningkatkan mutu pelayanan terhadap pelanggan yang menginap di hotel mereka agar meningkatkan rating hotel di *platform trip advisor*.
4. Membantu pelanggan menemukan hotel di *platform trip advisor* yang memberikan pelayanan yang memuaskan dan memiliki rating yang baik.