

BAB I

PENDAHULUAN

1.1. Latar Belakang

Pemerintah membutuhkan informasi tentang pariwisata untuk menentukan kebijakan-kebijakan terkait pengembangan sektor pariwisata. Salah satu parameter yang digunakan untuk membuat kebijakan adalah tingkat kunjungan wisatawan (Kementerian Pariwisata, 2016). Tingkat kunjungan wisatawan dipengaruhi oleh kepuasan wisatawan terhadap fasilitas yang ada di daerah tujuan wisata (Ni Nyoman Ayu Wiratini M, 2018).Kepuasan konsumen dapat didefinisikan sebagai: “Respon konsumen terhadap evaluasi diskrepansi/ketidaksesuaian yang dirasakan antara ekspektasi sebelumnya (atau beberapa norma kinerja lain) dan kinerja aktual dari produk sebagaimana yang dirasakan setelah pengkonsumsianya”. (Wilton, 1998)

Hotel merupakan salah satu sarana akomodasi utama untuk menjadi prioritas bagi wisatawan untuk menentukan tujuan wisatanya. Tingkat kepuasan wisatawan terkait akomodasi dapat dilihat dari *review* penggunaan hotel. *Review* penggunaan hotel dapat dilihat pada *feedback* dan *review* yang diberikan oleh wisatawan setelah menggunakan fasilitas hotel terkait. Sehingga, melalui informasi *feedback* dari wisatawan bisa didapatkan parameter tingkat kepuasan wisatawan selama penggunaan fasilitas hotel. Penelitian membuktikan secara signifikan bahwa loyalitas konsumen terhadap hotel dipengaruhi oleh kepuasan konsumen (Dyah Sugandini, 2018).

Biasanya wisatawan tidak memberikan *review* secara langsung terkait *feedback* ke sebuah hotel, melainkan menyalurkannya ke forum-forum diskusi, sosial media dan situs-situs terkait termasuk *booking online hotel*. Forum diskusi, sosial media dan *booking online hotel* tersebut banyak tersedia dalam bentuk *website*. *Feedback* atau *review* yang disampaikan melalui *website* berupa teks. Halaman sosial media sebuah *brand* telah menjadi instrumen penting yang memungkinkan konsumen berpartisipasi secara sukarela memberikan *feedback*

untuk memberikan masukan ide peningkatan dan berkolaborasi dengan konsumen yang lain (Carlson, 2018).

Dari uraian diatas, *feedback* yang ada pada *website* dapat menjadi sumber data yang dibutuhkan dalam evaluasi tingkat kepuasan wisatawan. Menurut beberapa penulis, *feedback* dan *review* yang ada di *website* merupakan data tidak terstruktur (Plamen Milev, 2017). Untuk mengolah data *review* para wisatawan, dibutuhkan data yang terstruktur. Oleh karena itu dibutuhkan sebuah sistem yang dapat mengambil data dari *website* dan menghasilkan data *review* dari wisatawan yang terstruktur.

Sistem *scraping* dan *crawling* digunakan untuk mengekstraksi data dari *website* dan menghasilkan data yang terstruktur dalam model data. Penggunaan *javascript* dan *ajax* pada sebuah *website* membuat akses data pada sebuah website tidak memerlukan *refresh* keseluruhan halaman web. Data pada *website* dapat ditampilkan dengan lebih interaktif. Untuk melakukan *crawling* pada *website* yang menggunakan *javascript* dan *ajax* diperlukan teknik tertentu sehingga sistem *scrawling* dapat berinteraksi dengan *ajax* dan proses *scraping* dapat mengambil semua data yang ada pada sebuah halaman web. Pengembangan teknik web *scraping* merupakan proses yang cukup kompleks (Plamen Milev, 2017). Sehingga dibutuhkan pendekatan teknik *scraping* yang disesuaikan dengan sumber *website*. Setelah mendapatkan data yang terstruktur, data tersebut disimpan dan dapat digunakan untuk proses lebih lanjut, misalnya proses statistik, proses sentimen analisis, dan lain-lain.

1.2. Rumusan masalah

Mengekstraksi data *review* hotel dari sebuah *website* sangat dibutuhkan untuk menggali informasi yang ada pada data dalam proses teks *mining*. *Website* merupakan dokumen teks *html* yang datanya tidak terstruktur. Data *review* hotel yang ada pada dokumen *html* juga bercampur dengan teks-teks lain yang bukan merupakan data. Selain itu akses data pada halaman web menggunakan *ajax* dimana perubahan halaman web dilakukan tanpa semua halaman web. Proses untuk

mengekstraksi dan mengolah informasi membutuhkan data yang terstruktur dalam sebuah model data.

Oleh karena itu, dibutuhkan sebuah sistem yang dapat melakukan ekstraksi data dari sebuah *website* berbasis *ajax*, melakukan *crawling* web berbasis *ajax* dan menghasilkan data terstruktur dalam sebuah model data.

1.3. Batasan masalah

Pengembangan sistem *scraping* dan *crawling* dibuat untuk mengekstraksi data dari *website*. Pengembangan diutamakan pada proses ekstraksi data untuk menghasilkan data terstruktur sesuai dengan model data yg telah ditentukan. Hal-hal yang tidak termasuk dalam pengembangan ini adalah:

1. Topologi jaringan dari sistem *scraping* ke *website* yang diproses
2. Proses pengolahan data hasil ekstraksi
3. Basis data yang digunakan untuk menyimpan hasil ekstraksi

1.4. Tujuan

Pengembangan sistem *scraping* dan *crawling* ini bertujuan untuk membangun sistem yang dapat melakukan proses *scraping* dan *crawling* web berbasis *ajax* untuk menghasilkan data *review* hotel terstruktur. Data hasil ekstraksi dapat disimpan kedalam database untuk proses lebih lanjut.

1.5. Manfaat

Manfaat dari dikembangkannya sistem *scraping* dan *crawling* ini adalah agar lebih mudah mendapatkan data terstruktur dari sebuah *website*. Mengekstraksi data menggunakan sistem ini sangat membantu agar lebih cepat mendapatkan data dibandingkan melakukan pengambilan data secara manual. Dengan tersedianya data yang dibutuhkan, pemerintah dapat menentukan kebijakan-kebijakan terkait pengembangan sektor pariwisata melalui data atau informasi yang dihasilkan dari proses ekstraksi ini.

Manfaat lain dari proses ekstraksi ini adalah untuk menghasilkan data yang dibutuhkan dari *website* untuk kepentingan perusahaan atau individu. Ketika suatu perusahaan akan membuat atau mengembangkan suatu perangkat atau aplikasi, dibutuhkan analisa mengenai apa yang akan dibuat. Analisa tersebut dapat berupa data yang didapat dari calon pengguna secara langsung, maupun data dari para pengguna yg ada pada perangkat atau aplikasi yang pernah dibuat. Data tersebut dibutuhkan untuk mengukur kebutuhan calon pengguna, melihat respon dari pengguna di aplikasi sebelumnya dan lain-lain. Dapat juga untuk melihat seberapa baik aplikasi yang sebelumnya dibuat, dalam hal ini bertujuan agar perusahaan dapat mengembangkan aplikasi menjadi lebih baik dan lebih nyaman digunakan oleh para pengguna.