

# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang Masalah**

World Wide Web saat ini menjadi fasilitas yang sangat populer dan interaktif dalam melakukan pertukaran informasi. Ratusan juta data terus bertambah dengan cepat ke internet seiring dengan jumlah pengguna yang juga bertambah. Dengan banyaknya informasi yang ada di dalam web, pengguna seringkali harus lebih lihai dalam melakukan pencarian data agar sesuai dengan keinginannya. Bagi sebagian pengguna ini adalah hal yang memusingkan dan menghabiskan banyak waktu. Disisi lain, pelaku e-business dan web marketer menginginkan agar produk dan layanannya dapat sampai kepada pengguna yang tepat. Dari kedua hal ini dapat dilihat bahwa mendapatkan keinginan pengguna atau sering disebut personalisasi menjadi hal yang sangat penting. Analisis dari pola pengguna ketika melakukan browsing, dapat memberikan informasi yang penting kepada penyedia layanan. Web personalisasi dapat digambarkan sebagai beberapa aksi yang dapat membangun web experience dari seorang pengguna berdasarkan ketertarikan dari pengguna (Vellingiri dan Pandian, 2011).

Industri pariwisata saat ini mulai memanfaatkan fasilitas World Wide Web. Gretzel dkk (2000) mengatakan internet adalah sumber informasi utama dalam domain pariwisata. Dengan semakin meningkatnya informasi dalam dunia internet mengenai tujuan pariwisata menjadikan para wisatawan tidak lagi bergantung pada agen perjalanan (Arbelaitz dkk, 2013). Para wisatawan lebih suka untuk melihat informasinya melalui internet, baik itu informasi iterinary nya maupun ulasan individual dari setiap tempatnya. Namun ketersediaan informasi yang terlalu banyak di berbagai sumber di internet seringkali memunculkan kebingungan kepada para wisatawan pemula. Informasi yang keluar seringkali kurang sesuai dengan keinginan wisatawan. Sebuah sistem rekomendasi dapat menjadi alat bantu untuk menyelesaikan masalah tersebut (Buhalis dan Law, 2008). Sistem rekomendasi juga dapat menyediakan informasi yang berguna untuk membantu pengguna dalam menentukan pilihannya. Untuk dapat melakukan hal tersebut sistem rekomendasi harus dapat mengidentifikasi keinginan pengguna berdasarkan pengguna lain yang memiliki preferensi yang sama (Fesenmair dkk, 2003).

Saat ini, teknik untuk membuat suatu sistem profil navigasi pengguna sering dilakukan dengan menganalisis server log file untuk mencari pola akses pengguna yang kemudian disebut dengan web usage mining. Informasi semantik yang didapat dari konten juga dapat menambah kesesuaian profil yang dihasilkan nantinya (Arbelaitz, 2012). Oleh karena itu, dalam penelitian ini telah dibuat suatu sistem dengan memanfaatkan web usage mining dan web content mining untuk dapat melakukan pemodelan profil navigasi pengguna dengan studi kasus pada website *gudegnet*. *Gudegnet* mempunyai lebih dari 7.000 konten dengan jumlah kunjungan yang cukup besar tiap harinya. Dengan studi kasus ini, diharapkan sistem dapat diimplementasikan juga kepada website lain yang mempunyai konten yang cukup besar seperti portal berita ataupun portal informasi lainnya.

## **1.2 Perumusan Masalah**

Dari uraian Latar Belakang, dapat dirumuskan permasalahan, yaitu bagaimana membuat sistem dalam pembuatan profil navigasi pengguna dengan melakukan kombinasi antara web usage mining dan web content mining.

## **1.3 Tujuan Penelitian**

Tujuan penelitian ini adalah mendapatkan suatu sistem dalam melakukan web usage mining dan web content mining untuk pembuatan profil navigasi pengguna website, yang nantinya oleh penyedia layanan website dapat digunakan sebagai alat untuk memodelkan para pengunjung websitenya.

## **1.4 Batasan Masalah**

Berdasarkan permasalahan yang telah dikemukakan, permasalahan dibatasi pada:

1. Studi kasus dilakukan di website *gudegnet*.
2. Analisis web content mining hanya dalam lingkup halaman yang memiliki konten artikel teks, tidak termasuk di dalamnya yaitu galeri, tentang kami, hubungi kami, halaman administrator, aturan, forum, dan peta situs.

## 1.5 Manfaat Penelitian

Manfaat yang ingin dicapai dari penelitian ini yaitu, memberikan alat kepada penyedia layanan website untuk melihat profil navigasi pengguna websitenya, terutama pada website yang mempunyai konten artikel yang besar. Profil navigasi ini dapat bermanfaat sebagai landasan untuk melakukan pengembangan website.

## 1.6 Keaslian Penelitian

Penelitian ini membahas tentang pembuatan sistem kombinasi web usage mining dengan web content mining untuk mencari profil navigasi pengguna website dengan menggunakan algoritma Latent Dirichlet Allocation (LDA), Self Organizing Map (SOM), dan PrefixSpan.

## 1.7 Metodologi Penelitian

Metode penelitian dalam mengembangkan sistem ini bisa dibagi menjadi beberapa tahap, yaitu:

### 1. Studi literatur

Studi literatur dilakukan dengan membaca buku-buku, jurnal-jurnal, sumber-sumber di Internet, serta memahami proses yang terkait dengan web usage mining dan web content mining.

### 2. Pengumpulan data

Pengumpulan data dilakukan di salah satu website pariwisata di Yogyakarta yaitu *gudegnet*. Data yang dikumpulkan adalah log akses ke website selama satu bulan serta konten artikel dari website.

### 3. Perancangan sistem

Pada tahap ini, dirancang alur sistem yang dibangun seperti dijelaskan secara lebih mendetail di Bab 4 dengan menggunakan Diagram Alir Data. Tampilan antar muka keseluruhan sistem juga dirancang pada tahap ini.

### 4. Implementasi

Pada tahap ini, hasil rancangan sistem dikembangkan menjadi perangkat lunak menggunakan bahasa pemrograman python serta beberapa tools bantuan.

#### 5. Pengujian sistem

Pada tahap ini, pengujian dilakukan terhadap sistem yang telah dibuat dengan sumber data dari *gudegnet*.

#### 6. Evaluasi sistem

Profil navigasi pengguna yang dihasilkan dievaluasi dengan menerapkan fitur sistem rekomendasi tautan pada sebuah artikel. Tautan yang direkomendasikan diberikan berdasarkan profil navigasi pengguna yang telah terbuat sebelumnya. Pengguna baru yang menavigasi sebuah artikel dilihat kecenderungannya untuk mengikuti rekomendasi tautan yang diberikan. Evaluasi terhadap profil navigasi pengguna yang sudah terbuat dilakukan dengan melihat precision, recall, dan F-Measure.

## 1.8 Sistematika Penulisan

### BAB I PENDAHULUAN

Bab ini berisi uraian mengenai latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, tinjauan pustaka, metodologi penelitian, dan sistematika penulisan.

### BAB II TINJAUAN PUSTAKA

Bab ini berisi tentang informasi hasil penelitian yang telah dilakukan sebelumnya, dimana penelitian tersebut memiliki keterkaitan dengan penelitian yang dilakukan.

### BAB III LANDASAN TEORI

Bab ini berisi Landasan Teori, yang merupakan pembahasan tentang teori dasar yang berkaitan dengan konsep dasar sistem, web usage mining, web content mining, pemodelan topik dengan LDA, Clustering dengan SOM, dan *Sequential Pattern Mining* dengan PrefixSpan.

### BAB IV ANALISIS DAN RANCANGAN SISTEM

Bab ini berisi uraian mengenai analisis dan kebutuhan perancangan sistem yang dibuat. Bab ini juga menguraikan tentang perancangan basis data, Diagram Alir Data, flowchart, dan antar muka (interface).

### BAB V IMPLEMENTASI

Bab ini berisi potongan kode program (sourcecode) dan interface program dari implementasi sistem yang telah dihasilkan.

### BAB VI HASIL DAN PEMBAHASAN

Bab ini berisi hasil akhir sistem yang dibangun berdasarkan hasil pengujian sistem dan disertai dengan penjelasan pada keluaran proses.

### BAB VII KESIMPULAN DAN SARAN

Bab ini berisi Kesimpulan dan saran, dalam hal ini ada beberapa kesimpulan dari proses pengembangan sistem, baik pada tahap analisis kebutuhan sistem, perancangan sistem, implementasi sistem, dan terutama pada analisis kinerja sistem serta saran-saran berdasarkan keterbatasan dan kekurangan yang ditemukan pada penelitian ini.