

# BAB I. PENDAHULUAN

## 1.1. Latar Belakang

Pada saat ini informasi teknologi merupakan suatu bidang yang dapat menarik perhatian para pakar dan peneliti. Peran informasi teknologi dalam kehidupan sehari-hari sangatlah signifikan, sehingga perubahan sekecil apapun dalam pengolahan informasi ini akan menjadi dampak bagi masyarakat sekitar. Oleh karena itu, penelitian dan pembelajaran dalam bidang ini sangatlah diperlukan. Di Indonesia terdapat 72.7 juta pengguna *internet* dan 72 juta akun media sosial yang aktif dimana 62 juta diantaranya adalah pengguna *mobile phones* (Kemp, 2016). Secara garis besar, dari total penduduk yang berjumlah 255 juta orang, 30% di antaranya sudah terhubung dengan dunia maya. Angka ini menyimpulkan bahwa semakin banyak orang di Indonesia bergantung terhadap *internet* untuk mendapatkan informasi dan berinteraksi. Perusahaan *start-up* seperti Tokopedia, Traveloka dan Bukalapak melihat peluang dan menjadi jawaban dari kebutuhan pengguna *internet* tersebut. Dengan banyaknya informasi yang menyebar, perusahaan-perusahaan ini memanfaatkan informasi dari pengguna jasa layanan mereka. Mereka memproses setiap *browsing history*, *purchase history* dan bahkan informasi dari media sosial mereka untuk merangkai strategi *marketing*. Dari strategi yang mereka dapatkan, akurasi dalam melakukan *advertising* untuk mencapai target akan meningkat. Hal ini dapat dirasakan juga dari sisi pengguna, dimana mereka akan lebih mudah untuk mencari dan mendapatkan informasi yang relevan dengan apa yang mereka cari.

Perusahaan *start-up* yang bernama Yelp berasal dari Amerika dan didirikan pada tahun 2004. Tujuan dari perusahaan ini adalah untuk membantu masyarakat untuk mencari tempat usaha di daerah sekitar mereka berada. Menurut Yelp.com pada saat ini pengguna yang menggunakan *Yelp mobile app* mencapai rata-rata 21 juta per bulannya dan terdapat 69 juta pengguna Yelp pada kuartal pertama tahun 2016. Para pengguna Yelp inilah yang mempopulasi data dengan ulasan di berbagai jenis usaha. Pada kuartal pertama tahun 2016, terdapat 102 juta ulasan dari seluruh pengguna aplikasi dan *web* Yelp.

Maka terciptalah suatu kebutuhan untuk mengetahui apa yang akan dicari dan relevan pada kebutuhan *users* dari aplikasi-aplikasi ini. Oleh karena itu perlu adanya prediksi untuk mengetahui hubungan antara *user to item* dan *item to item*. Untuk mendapatkan informasi *user* dan *item*, kita perlu untuk mendapatkan informasi tentang perilaku *user* di lingkungannya, pada kasus ini terdapat di aplikasi *Yelp!* Salah satu cara yang dapat digunakan untuk mendapatkan informasi adalah untuk merekam *browsing history* atau mendapatkan hasil dari *ratings* yang diberikan oleh *users*. Itulah sebabnya kenapa informasi ini sangatlah berharga, karena semakin banyak pengguna yang menggunakan Yelp maka semakin banyak informasi yang akan di dapatkan untuk dianalisa dan digunakan untuk mengklasifikasikan kebutuhan pengguna berdasarkan dari karakteristik mereka.

Seperti yang diketahui, rekomendasi dilakukan untuk mengetahui *item* yang bisa direkomendasikan kepada pengguna lainnya. Hasil rekomendasi sangat berguna untuk menentukan tindakan apa yang harus dilakukan sekarang untuk melayani pengguna dengan karakteristik tertentu. Saat ini penelitian mengenai

sistem rekomendasi telah banyak dilakukan di banyak domain, seperti (Linden, Smith, & York, 2003) memprediksi pengguna Amazon.com , (Davidson, Liebald, Liu, Nandy, & Van Vleet, 2010) menelusuri dan merancang video *recommendation system*, dan (Resnick, Iacovou, Suchak, Bergstrom, & Riedl, GroupLens: An Open Architecture for Collaborative Filtering of Netnews, 1994) dalam membangun sebuah sistem untuk melakukan *Collaborative Filtering* yang di namakan *Group Lens*.

Banyak metode yang digunakan untuk mendapatkan hasil rekomendasi yang akurat. Salah satunya adalah metode *Collaborative Filtering*. Metode ini sering digunakan dalam pembangunan sebuah sistem rekomendasi. Dalam *Collaborative Filtering* terdapat dua jenis *filtering* yang umumnya digunakan, yaitu *user-based collaborative filtering* dan *item-based collaborative filtering*. Pada *User-based collaborative filtering* prediksi berdasarkan kepada pengguna lainnya yang memiliki kemiripan properti, sebagai contoh sebagian besar orang di Jakarta menghabiskan waktunya untuk pergi ke *shopping mall*, sedangkan pengguna yang berasal dari Malang memilih untuk tinggal diam di sekitar rumah. Pada *item-based collaborative filtering*, prediksi rekomendasi akan di dasarkan kepada *item* dengan properti yang sama, sebagai contoh orang yang membeli *iPhone* memiliki besar kemungkinan untuk membeli sebuah *powerbank* juga.

Penelitian ini akan mengimplementasikan metode *item-based collaborative filtering* terhadap *dataset* yang didapatkan dari Yelp Metode *item-based collaborative filtering* akan memberikan informasi terhadap suatu *item* (tempat) yang direkomendasikan karena populer di daerah kota tersebut. Jika pengguna telah memberikan *rating* terhadap suatu tempat di suatu kota, maka sistem ini

akan memberikan rekomendasi tempat-tempat yang diprediksi akan dikunjungi juga setelah mengunjungi tempat tersebut. Dengan informasi ini, akan kita dapatkan informasi demografik tentang tempat-tempat yang umumnya dikunjungi pada suatu kota.

## 1.2. Perumusan Masalah

Permasalahan yang diangkat pada studi kasus ini adalah bagaimana menggunakan tempat-tempat yang dapat direkomendasikan berdasarkan data Yelp di Pittsburgh, Phoenix dan Las Vegas.

## 1.3. Tujuan Penelitian

Tujuan studi kasus ini adalah:

- 1) Menentukan tempat-tempat yang dapat direkomendasikan berdasarkan metode *item-based collaborative filtering* dan *user-based collaborative filtering* pada Yelp dataset.
- 2) Analisis keakuratan hasil prediksi *item-based collaborative filtering* dan *user-based collaborative filtering*.
- 3) Memberikan informasi demografik dari hasil *item-based collaborative filtering* dan *user-based collaborative filtering*.

## 1.4. Manfaat Penelitian

Manfaat penelitian ini adalah:

- 1) Menyediakan rekomendasi yang dapat digunakan para pendatang mengenai tempat-tempat yang umumnya dikunjungi di kota Pittsburgh, Phoenix dan Las Vegas.

- 2) Menyediakan informasi spasial untuk para wiraswasta untuk dapat memprediksi lokasi yang strategis di kota Pittsburgh, Phoenix dan Las Vegas.

## 1.5. Ruang Lingkup

Ruang lingkup penelitian ini adalah sebagai berikut:

- 1) Menyaring *dataset* yang diperoleh dari *Yelp* sehingga *review* dan lokasi bisnis yang digunakan hanya yang berada dari kota Pittsburgh, Phoenix dan Las Vegas.
- 2) Membangun sistem rekomendasi dengan menggunakan metode *item-based collaborative filtering* dengan bahasa Java dan *tool* Eclipse .
- 3) Menganalisis keakuratan hasil prediksi *item-based collaborative filtering* dengan menggunakan *Root Mean Square Error* (RMSE)