

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Mencari berita yang sesuai dan diinginkan belakangan ini memerlukan waktu yang lama. Dengan adanya pertumbuhan pada industri media dan internet, pengguna internet sering kali tidak dapat mengakses informasi yang mereka butuhkan dengan cepat dan efisien [1]. Dalam waktu yang sama pertumbuhan tersebut malah menyebabkan masalah kelebihan informasi atau *information overload*. Kelebihan informasi merupakan fenomena dimana kita kesulitan dalam membuat keputusan secara efektif ketika kita memiliki terlalu banyak informasi dan membuat kita sulit untuk mencari dan mengambil apa yang kita inginkan. Informasi yang berlebihan dapat menghambat kita dalam mengakses ke informasi yang kita cari, sehingga mengurangi efisiensi penggunaan informasi [2], [3].

Sistem rekomendasi dapat membantu kita untuk memecahkan masalah *information overload*, dengan cara menyaring aliran informasi yang masuk sesuai dengan preferensi pembaca sehingga menghasilkan *personalized news content*. Konsep fundamental dari *personalized news content* adalah memperlakukan setiap pembaca berita sebagai entitas yang unik dan memberikan berita berdasarkan dengan karakteristik dan preferensi pembaca, *personalized news content* diharapkan dapat lebih diingat, disukai dan lebih persuasif [4].

Tantangan dalam menerapkan *personalized news content* adalah pengguna harus mendapatkan campuran berita yang berhubungan dengan kesukaan mereka yang memberikan informasi dan hiburan bagi pengguna. Dalam membangun *personalized news content* diperlukan data *log reading history* pengguna untuk

menilai tingkat kesukaan pengguna terhadap ketertarikan berita yang mereka inginkan.

Dengan menggunakan log riwayat artikel berita yang telah dibaca pengguna, log dapat dimanfaatkan untuk meningkatkan kepuasan pelanggan dengan menawarkan *personalized news content*. Hal ini berguna untuk meningkatkan frekuensi kunjungan, yang akan meningkatkan keuntungan dari iklan, meningkatkan *user engagement* dan *retention*, sehingga meningkatkan dan mempertahankan jumlah langganan [5], [6], [7].

*Content-based filtering* atau CBF adalah suatu teknik *filtering* yang menggunakan deskripsi sebuah *item* dan preferensi pengguna berdasarkan log riwayat pengguna [8]. Untuk merekomendasikan suatu *item* yang mempunyai deskripsi atau berisi teks seperti artikel berita, CBF adalah metode yang paling baik untuk digunakan dibandingkan metode lain seperti *Collaborative Filtering* [2]. CBF tidak memerlukan data dari pengguna lain untuk merekomendasi suatu item sehingga CBF tidak dipengaruhi oleh *cold start problem*. Sistem tetap dapat merekomendasikan suatu artikel meskipun artikel yang direkomendasikan itu baru dan pembacanya sedikit. Untuk menghasilkan rekomendasi yang berarti atau relevan, CBF menggunakan berbagai jenis model yang dapat membantu untuk menghasilkan rekomendasi yang relevan.

TF-IDF adalah metode yang merupakan salah satu Vector Space Model dan digunakan untuk menentukan seberapa penting suatu kata dalam dokumen teks [9]. TF-IDF sering digunakan untuk faktor pembobotan dalam information extraction. Survey menyatakan bahwa banyak penelitian lain yang menggunakan CBF juga

menggunakan TF-IDF untuk perhitungan bobot [10], [11] guna menghasilkan rekomendasi yang relevan. Oleh karena itu, penulis akan menggunakan CBF dan TF-IDF dalam penelitian ini untuk merancang sistem rekomendasi.

Pengujian dilakukan dengan menggunakan metode *precision@k*, *recall@k*, dan *f-measure* untuk mengetahui apakah sistem dapat merekomendasikan artikel berita yang sesuai dengan aktivitas pengguna Kompas.id. [12]

## **1.2 Identifikasi Masalah**

Berdasarkan latar belakang telah disebutkan sebelumnya, rumusan masalah penelitian ini adalah:

1. Apakah sistem rekomendasi dapat menampilkan konten secara akurat?

## **1.3 Batasan Penelitian**

Batasan masalah penelitian yang ditetapkan oleh peneliti adalah sebagai berikut:

1. Penulis akan menggunakan artikel berita dan data pengguna dari situs Kompas.id
2. Penulis hanya akan mengevaluasi akurasi rekomendasi yang dihasilkan oleh sistem.

#### **1.4 Tujuan Penelitian**

Tujuan dari penelitian ini adalah membuat sistem rekomendasi artikel berita menggunakan metode *content-based filtering* dan algoritma TF-IDF yang dapat merekomendasikan artikel yang mirip dengan artikel yang telah dibaca oleh pengguna situs Kompas.Id secara akurat.

#### **1.5 Manfaat Penelitian**

Manfaat dari penelitian ini adalah sebagai berikut:

1. Mempermudah pengguna untuk mencari dan mengambil konten yang diinginkan, sehingga mengurangi efek *information overload*.
2. Meningkatkan tingkat kepuasan pengguna situs Kompas.Id