

# BAB I PENDAHULUAN

## 1.1. Latar Belakang Penelitian

Data dan informasi merupakan aset yang sangat berharga, terutama di era digital yang dipenuhi oleh jutaan bahkan miliaran konten baru setiap harinya. Namun, di saat yang bersamaan, menemukan informasi yang sesuai dan dibutuhkan menjadi aktivitas yang melelahkan serta memakan banyak waktu. Permasalahan ini juga dapat memicu dampak lain, seperti mudahnya penyebaran hoaks dan menurunnya minat baca akibat banyaknya informasi yang tidak relevan. Oleh karena itu, dibutuhkan sistem yang mampu menyaring dan merekomendasikan informasi secara otomatis dan relevan.

Menurut Mondy et al. (2019), sistem rekomendasi merupakan perangkat lunak atau sistem yang mampu memberikan saran item kepada pengguna [1]. Item yang disarankan tersebut dapat berupa produk, gambar, atau teks yang memiliki atribut tertentu, dan dianggap berguna bagi pengguna yang bersangkutan.

Pada penelitian ini, objek yang direkomendasikan adalah bahan bacaan yang disesuaikan dengan rating pengguna. Untuk menjalankan tugas tersebut, metode *Collaborative Filtering*, khususnya *Model-Based Collaborative Filtering*, dipilih karena mampu menyarankan item berdasarkan kesamaan rating antar pengguna dan tidak memerlukan data yang terlalu detail.

*Collaborative Filtering* merupakan metode yang memberikan rekomendasi berdasarkan interaksi atau kesamaan minat antara pengguna [1]. Sistem ini bekerja dengan cara menemukan pola kemiripan antara satu pengguna dengan pengguna lainnya, kemudian membandingkan rating mereka untuk memprediksi nilai atau minat terhadap suatu item.

Salah satu kelemahan utama dari metode ini adalah ketergantungannya terhadap data interaksi pengguna. Ketika sistem tidak memiliki informasi mengenai pengguna baru, maka sistem tidak dapat memberikan rekomendasi. Masalah ini dikenal sebagai *cold-start problem* [2].

Untuk mengatasi permasalahan tersebut, dalam penelitian ini akan ditambahkan fitur pendukung berupa sistem klasifikasi teks. Menurut Prakoso et al.

(2017), sistem klasifikasi teks adalah suatu sistem yang berfungsi untuk mendefinisikan satu atau lebih kategori terhadap dokumen dalam bahasa natural [3]. Pada penelitian ini, dokumen yang diklasifikasikan berupa judul bacaan yang tersedia, yang kemudian dikelompokkan ke dalam kategori tertentu. Proses klasifikasi otomatis akan dibangun menggunakan algoritma *Naive Bayes*.

Algoritma *Naive Bayes* merupakan algoritma klasifikasi yang didasarkan pada prinsip teori probabilitas, dan banyak digunakan untuk menyelesaikan permasalahan statistik [4]. Algoritma ini sangat efektif dalam menangani data berbentuk teks, sehingga sesuai digunakan untuk mengklasifikasikan judul bacaan berdasarkan kategori yang relevan. Dengan demikian, pembaca dapat menemukan bacaan yang sesuai dengan minat mereka secara lebih mudah dan efisien.

Selain keefektifannya dalam klasifikasi teks, algoritma *Naive Bayes* memiliki beberapa kelebihan lainnya, antara lain:

1. Algoritma ini memiliki kecepatan eksekusi yang baik, sehingga dapat mempersingkat waktu.
2. Algoritma ini kompatibel untuk melakukan prediksi multi-kelas.
3. Algoritma ini membutuhkan lebih sedikit data dibandingkan algoritma setipe.
4. Algoritma ini kompatibel untuk variabel *input* kategoris dibandingkan variabel numerik [5].

Dalam penelitian ini, algoritma *Naive Bayes* akan digunakan untuk mengelompokkan bacaan ke dalam beberapa kategori tertentu. Dengan pengelompokan tersebut, setiap bacaan akan diklasifikasikan ke dalam kategori yang paling sesuai berdasarkan kata kunci yang terdapat dalam judul, sehingga dapat meminimalkan kesalahan klasifikasi seperti penempatan kategori yang tidak tepat.

Berdasarkan uraian di atas, dapat disimpulkan bahwa implementasi sistem rekomendasi dengan metode *Collaborative Filtering*, serta sistem klasifikasi teks dengan algoritma *Naive Bayes*, merupakan perpaduan yang saling mendukung dalam menyajikan bacaan yang relevan. Penelitian ini akan membahas perancangan dan evaluasi kedua pendekatan tersebut dalam konteks pengelompokan dan penyajian bacaan digital secara lebih terarah dan efisien.

## 1.2. Rumusan Masalah Penelitian

Dari latar belakang yang telah dipaparkan, berikut merupakan rumusan masalah dari penelitian ini:

1. Bagaimana penggunaan algoritma *Naive Bayes* dalam mengklasifikasi judul bacaan berdasarkan kategori yang tersedia?
2. Bagaimana penggunaan metode *Collaborative Filtering* dalam merekomendasikan bacaan berdasarkan kesamaan rating pengguna?
3. Bagaimana akurasi dari *Naive Bayes* dalam melakukan klasifikasi dan *Collaborative Filtering* dalam melakukan rekomendasi ketika diuji menggunakan metode evaluasi kuantitatif?

## 1.3. Tujuan Penelitian

Dari latar rumusan masalah yang telah dipaparkan, berikut merupakan tujuan dari penelitian ini:

1. Membangun sistem klasifikasi judul bacaan menggunakan algoritma *Naive Bayes* agar bacaan dapat dikelompokkan ke dalam kategori yang sesuai.
2. Menerapkan metode *Collaborative Filtering* untuk memberikan rekomendasi bacaan berdasarkan interaksi antar pengguna.
3. Mengevaluasi performa sistem klasifikasi dan sistem rekomendasi yang dihasilkan dengan menggunakan metrik evaluasi seperti *Accuracy*, MAE dan RMSE.

## 1.4. Batasan Masalah Penelitian

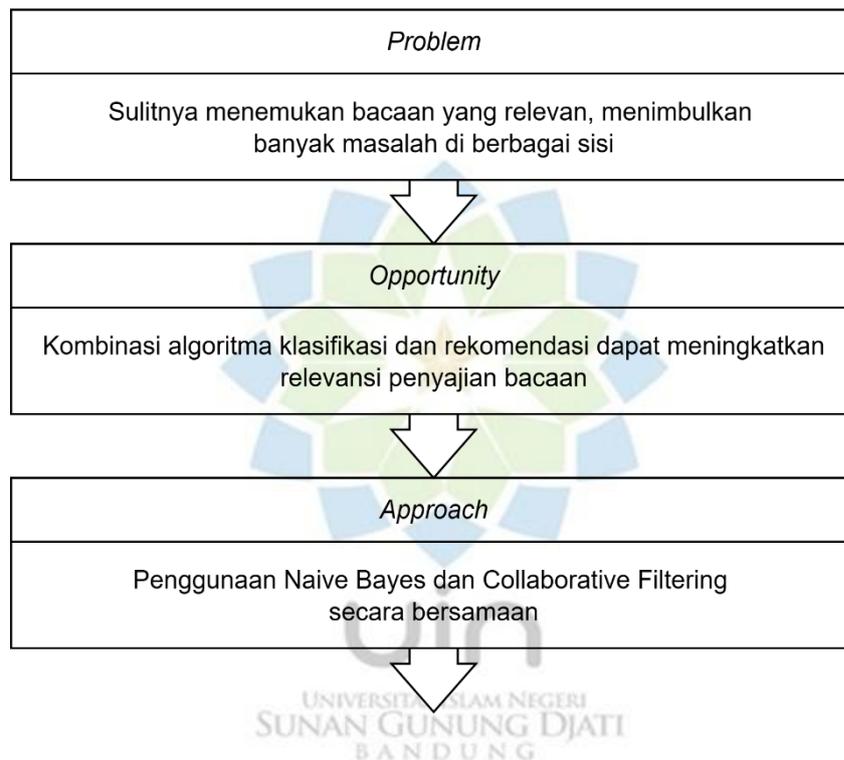
Batasan masalah dimaksudkan agar penelitian ini lebih terarah dan mencapai tujuan yang diharapkan dengan efisien. Berikut merupakan batasan masalah dari penelitian ini:

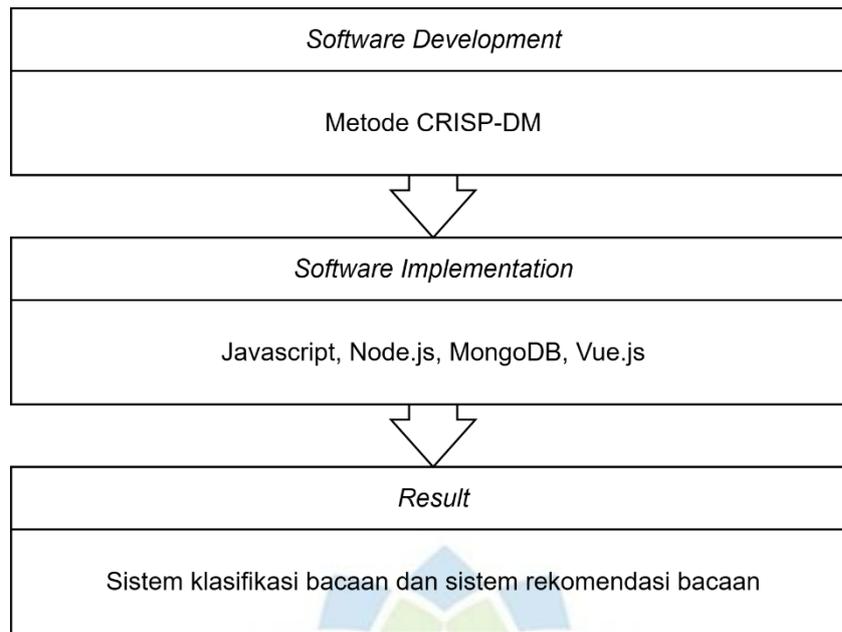
1. Penelitian ini menggunakan *Naive Bayes* untuk mengklasifikasikan judul bacaan, dan *Collaborative Filtering* untuk memberikan rekomendasi bacaan berdasarkan rating pengguna.
2. Sistem hanya memproses data dalam bentuk judul bacaan, kategori bacaan dan rating yang diberikan pengguna terhadap bacaan.
3. Klasifikasi bacaan ditentukan berdasarkan judul bacaan dan kategori terkait, sedangkan rekomendasi bacaan dihasilkan berdasarkan rating yang diberikan pengguna terhadap bacaan.

4. Evaluasi sistem klasifikasi dilakukan menggunakan metrik akurasi, sedangkan sistem rekomendasi dievaluasi menggunakan metrik MAE (*Mean Absolute Error*) dan RMSE (*Root Mean Squared Error*).
5. Penelitian ini tidak memperhitungkan waktu pemrosesan atau performa kecepatan dalam melakukan klasifikasi dan rekomendasi.

### 1.5. Kerangka Pemikiran Penelitian

Kerangka kerja tugas akhir ini dapat dilihat pada gambar berikut :





Gambar 1.1 Kerangka Pemikiran

Sebagaimana ditunjukkan pada Gambar 1.1, penelitian ini berawal dari kebutuhan akan sistem yang dapat mempermudah penyajian informasi, khususnya bahan bacaan. Kebutuhan ini membuka peluang untuk merancang sistem klasifikasi dan sistem rekomendasi yang dapat diintegrasikan ke dalam platform membaca digital. Untuk memenuhi kebutuhan tersebut, digunakan algoritma *Naive Bayes* yang efektif dalam melakukan klasifikasi teks, serta metode *Collaborative Filtering* yang mampu merekomendasikan bacaan berdasarkan pola kesamaan preferensi antar pengguna.

Penelitian ini menggunakan pendekatan CRISP-DM (*Cross Industry Standard Process for Data Mining*), yaitu metode yang umum digunakan dalam proyek berbasis data. Sistem dikembangkan menggunakan bahasa pemrograman JavaScript dengan Node.js sebagai *backend*, MongoDB sebagai basis data, dan Vue.js sebagai *framework* antarmuka pengguna. Dengan rancangan tersebut, diharapkan sistem klasifikasi dan rekomendasi dapat berjalan sesuai dengan tujuan yang telah ditetapkan.

### 1.6. Sistematika Penulisan

Seluruh proses dan hasil penelitian yang diperoleh akan didokumentasikan secara komprehensif ke dalam bentuk laporan. Struktur penulisan dari laporan

tersebut akan mengacu pada lima bab, yang masing-masingnya memuat konten sebagai berikut:

### **Bagian Awal Skripsi**

Merupakan bagian yang memuat halaman judul, abstrak, berbagai daftar konten, serta berbagai berkas kelengkapan laporan seperti persetujuan, pengesahan, keaslian skripsi, dan sebagainya.

### **Bagian Utama Skripsi**

#### **BAB I           PENDAHULUAN**

Merupakan bab yang berisikan latar belakang dari penelitian terkait, rumusan masalah dan tujuan yang ingin dicapai, serta beberapa hal terkait metodologi, kerangka penelitian, dan cara penulisan laporan.

#### **BAB II           KAJIAN LITERATUR**

Merupakan bab yang berisikan kumpulan konsep, teori, dan penelitian terkait yang dapat dijadikan referensi dalam penelitian. Bagian ini akan membandingkan penelitian yang sedang dilakukan dengan penelitian terkait, khususnya penelitian yang berkaitan dengan sistem rekomendasi.

#### **BAB III          METODOLOGI PENELITIAN**

Merupakan bab yang berisikan penjelasan, analisis, dan rancangan sistem yang akan dikembangkan. Bagian ini akan memuat analisis sistem, kebutuhan perangkat keras, kebutuhan perangkat lunak, serta perancangan arsitektur sistem.

#### **BAB IV          HASIL DAN PEMBAHASAN**

Merupakan bab yang memaparkan hasil dan melakukan pembahasan dari sistem yang telah dikembangkan.

#### **BAB V           SIMPULAN DAN SARAN**

Merupakan bab yang merangkum seluruh penelitian dengan kesimpulan dan saran.