

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pada masa pandemi covid-19 ini kita dituntut untuk melakukan pembatasan kegiatan di luar rumah, sehingga banyak kegiatan yang mulai bergeser dilakukan dari rumah masing – masing mulai dari sekolah hingga berkerja. Untuk menunjang kelancaran kegiatan tersebut maka diperlukan *device* yang sesuai dengan kebutuhan. Salah satu *device* yang bisa menunjang produktifitas baik untuk berkerja maupun kegiatan belajar mengajar adalah laptop. Permintaan laptop di Indonesia pada masa pandemi meningkat, pada kuartal II-2020, terjadi kenaikan permintaan komputer dan laptop sebesar 18,6% dengan jumlah pengiriman 38,6 juta unit [1]. Diperkirakan permintaan produk laptop akan terus mengalami peningkatan selama masa pandemic dari berbagai jenis laptop mulai untuk kebutuhan kantor hingga *gaming*.

Dikarenakan jenis laptop yang dirilis hingga saat ini sangat beragam maka diperlukan sistem rekomendasi yang berguna untuk membantu calon pembeli menemukan rekomendasi laptop yang sesuai dengan keinginan calon pembeli. Rekomendasi yang diberikan oleh sistem berdasarkan kemiripan spesifikasi dan informasi yang melekat pada laptop yang dipilih oleh calon konsumen.

Metode rekomendasi yang paling sering digunakan adalah *Content Based Filtering (CBF)*, *Collaborative Filtering (CF)*, *Knowledge Based Recommendation* dan *Demographic Recommendation*. Penulis menggunakan metode *Content Based Filtering* dikarenakan *CBF* dapat memberikan rekomendasi baru kepada calon pembeli berdasarkan kemiripan *feature* seperti

deskripsi dan keterangan lainnya yang dikandung oleh *item* sebelumnya yang disukai *user* tersebut [2].

Langkah awal dalam menentukan rekomendasi adalah mencari nilai *TF-IDF*. *TF* (*Term Frequency*) adalah jumlah berapa kali suatu kata muncul pada sebuah dokumen [3]. Untuk mencari *TF* sebuah kata rumusnya adalah banyaknya kata itu muncul di dalam dokumen dibagi dengan jumlah total *terms* di suatu dokumen [3]. Sedangkan *IDF* (*Inverse Document Frequency*) dari sebuah kata adalah ukuran seberapa signifikan kata tersebut di dalam korpus [4]. *IDF* dari sebuah kata dapat dicari dengan mencari log jumlah dari dokumen dibagi dengan jumlah dokumen dengan kata yang di cari di dalamnya [4].

Pada penelitian ini untuk melakukan pembobotan *term* dalam *corpus* menggunakan konsep *TF-IDF*. *TF-IDF* adalah salah satu bagian dari *Natural Language Processing (NLP)* yang digunakan untuk ekstraksi data, misalnya untuk menghitung frekuensi kemunculan sebuah kata di dalam dokumen yang kemudian memberikan pembobotan pada setiap kata [5].

Untuk mencari kemiripan antara kata kunci dengan dokumen menggunakan algoritma *Cosine Similarity*, yaitu adalah ukuran kesamaan antara dua buah *vector* dalam sebuah ruang dimensi yang didapat dari nilai cosinus sudut dari perkalian dua buah *vector* yang dibandingkan [5], hasil yang semakin mendekati angka 1 maka semakin mirip.

Dengan adanya sistem rekomendasi ini diharapkan calon pembeli dapat menemukan laptop yang sesuai dengan ketertarikan calon pembeli. Penggunaan metode *Content Based Filtering* diharapkan akan memberikan rekomendasi yang tepat kepada calon pembeli.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya maka dapat dirumuskan sebuah permasalahan sebagai berikut:

1. Bagaimana langkah – langkah menerapkan *Content Based Filtering* dalam sebuah sistem rekomendasi?
2. Bagaimana mengimplementasikan metode *content-based filtering* pada sebuah aplikasi *web*?
3. Bagaimana melakukan evaluasi hasil rekomendasi yang telah diimplementasikan?

## 1.3 Batasan Masalah

Untuk menghindari pembahasan yang melebar maka batasan masalah dalam penelitian ini antara lain:

- a. Metode rekomendasi menggunakan *Content Based Filtering*.
- b. Algoritma yang digunakan adalah *TF-IDF* dan *Cosine Similarity*
- c. *Dataset* menggunakan data laptop beserta spesifikasinya dari *website* sumber terbuka (*Kaggle.com*).
- d. Pembuatan aplikasi menggunakan framework *Flask*.
- e. Kode program ditulis menggunakan bahasa pemrograman *Python 3*.
- f. Metode pengujian menggunakan *NDCG (Normalized Discounted Cumulative Gain)*

## 1.4 Maksud dan Tujuan Penelitian

Maksud dan tujuan dari penelitian ini adalah membangun aplikasi sistem rekomendasi laptop berbasis *website* dan menerapkan metode *Content Based Filtering* dalam sistem rekomendasi.

## 1.5 Manfaat Penelitian

Dengan adanya penelitian ini, manfaat yang diharapkan adalah sebagai berikut:

1. Pengguna bisa mendapatkan rekomendasi laptop untuk di beli yang relevan dengan kebutuhan pembeli.
2. Sistem rekomendasi yang sudah jadi bisa menjadi referensi untuk membuat sistem rekomendasi lainnya selain laptop, karena secara mendasar sistem rekomendasi digunakan untuk membantu pengguna menemukan sesuatu yang mungkin diperlukan.

## 1.6 Metode Penelitian

Metodologi yang digunakan dalam penyusunan Skripsi ini adalah sebagai berikut:

### 1.6.1 Metode Pengumpulan Data

Data yang digunakan dalam penelitian ini diambil dari <https://www.kaggle.com/ionaskel/laptop-prices>. Pada data tersebut berisi daftar laptop dari berbagai perusahaan beserta spesifikasinya. Ada 12 kolom dalam data laptop ini yaitu: *Company, Product, TypeName, Inches, ScreenResolution, Cpu, Ram, Memory, Gpu, OpSys, Weight, Price\_euros*, harga dalam euro akan dikonversi ke dalam rupiah.

### 1.6.2 Metode Analisis

Pada tahap ini dilakukan analisis terhadap data yang akan diteliti, menentukan item yang dapat dijadikan item rekomendasi dan permasalahan algoritma yang digunakan.

### 1.6.3 Metode Perancangan

Pada tahap ini ada beberapa tahapan yang dilakukan, yang pertama menyiapkan data. Yang kedua mencari *TF-IDF* untuk melakukan pembobotan pada *term* dan dokumen, kemudian dilakukan *scoring* kemiripan menggunakan *Cosine Similarity*. Yang terakhir mengimplementasikan dalam sebuah aplikasi web menggunakan bahasa pemrograman *python* menggunakan *framework flask*.

### 1.6.4 Metode Pengujian

Pengujian yang dilakukan terhadap aplikasi web menggunakan metode *Black Box Testing* yaitu untuk mengetahui apakah aplikasi yang dibuat sudah berjalan sesuai dengan yang diharapkan. Sedangkan untuk mengevaluasi relevansi rekomendasi yang diberikan menggunakan metode *NDCG (Normalized Discounted Cumulative Gain)*.

### 1.7 Sistematika Penulisan

Agar penulisan laporan lebih terstruktur, sistematika penulisan dibagi menjadi beberapa bagian seperti berikut:

#### **BAB I PENDAHULUAN**

Pendahuluan ini menjelaskan tentang latar belakang, tujuan, rumusan masalah, batasan masalah, dan sistematika penulisan laporan penelitian.

#### **BAB II LANDASAN TEORI**

Bab ini menjelaskan tentang teori penunjang dan referensi berupa buku, jurnal, dan laporan penelitian.

#### **BAB III METODE PENELITIAN**

Bab ini menjelaskan tentang penjelasan mengenai desain penelitian, metode yang digunakan dalam penelitian serta *tools* apa saja yang digunakan untuk menunjang penelitian.

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

Bab ini menjelaskan tentang penjabaran penelitian dan pembahasan hasil penelitian.

#### **BAB V PENUTUP**

Bab ini menjelaskan tentang kesimpulan dari hasil akhir penelitian.

#### **DAFTAR PUSTAKA**

Memuat keterangan dari buku-buku dan *literature* lain yang menjadi acuan dalam penyusunan laporan penelitian ini.

