

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di era modern ini, perkembangan teknologi berkembang dengan pesat, sehingga membuat kebutuhan manusia hampir semuanya dilakukan secara digital. Salah satunya adalah platform e-commerce yang membuat kebutuhan akan belanja yang sebelumnya dilakukan offline sekarang ini semua belanja bisa dilakukan dengan online. Dengan banyaknya pengunjung atau pengguna platform e-commerce ini membuat transfer data di platform tersebut tumbuh dengan signifikan. Data merupakan kunci dalam kegiatan manusia di internet atau dalam belanja online. Dengan pertumbuhan data yang sangat besar ini membuat platform tersebut dapat memberikan informasi yang sesuai dengan kebutuhan pengguna, hal ini membuat pengguna dapat dengan mudah dalam mengaksesnya, sehingga menyebabkan pengguna internet atau platform e-commerce meningkat setiap harinya. Hal ini harus dapat diselesaikan dengan cepat dan efektif. Dengan hal ini mesin rekomendasi dapat digunakan untuk mempermudah dalam melakukan pencarian informasi. Berbagai produk yang ada pada e-commerce seperti pakaian, elektronik, kendaraan, furniture dan lainnya mengandalkan produknya menggunakan visualisasi untuk menarik pembeli. Dari produk tersebut terdapat jutaan data gambar. Berdasarkan data-data gambar tersebut, menampilkan informasi yang diinginkan oleh pelanggan adalah masalah yang tidak

mudah diselesaikan[1]. Banyak peneliti yang telah menemukan untuk mengatasi masalah ini tetapi masih belum dapat memuaskan pengguna atau pelanggan, karena pendekatan berbasis teks ini harus memasukkan deskripsi dari produk, yang tidak bisa menggambarkan produk sepenuhnya[2][3]. Beberapa pendekatan telah dilakukan oleh peneliti sebelumnya [4] dengan mengusulkan sistem content-based similar image retrieval menggunakan Particle Swarm Optimization (PSO) dan K-means. Algoritma PSO adalah algoritma yang berbasis populasi yang berfungsi untuk mengeksploitasi dalam pencarian individu yaitu dengan cara mengekstrak berbagai citra warna seperti histogram. Selain itu, peneliti lain [5] menggunakan 3 tahap dalam sistem rekomendasi berbasis konten. Langkah pertama dengan menghapus latar belakang dari gambar. Kedua, melakukan representasi model gambar dan dilanjutkan dengan mengekstrak dan menormalisasikan fitur. Terakhir adalah proses menghitung bobot setiap fitur berdasarkan jejak pengguna.

Sistem rekomendasi adalah suatu sistem yang digunakan user atau pengguna untuk mencari informasi mengenai produk yang diinginkan. Sistem ini bekerja dengan menggunakan informasi pengguna yang telah diakses. Tujuan dari sistem ini untuk meningkatkan penjualan produk bagi penyedia platform. Sedangkan bagi pengguna dapat memudahkan dalam menemukan informasi yang mereka inginkan dengan cepat dan efektif. Terdapat beberapa metode yang ada dalam sistem rekomendasi yaitu collaborative filtering, content-based filtering dan hybrid filtering[6]. Content-based filtering merupakan salah satu metode yang mudah diimplementasikan dalam sistem

rekomendasi. Metode ini nantinya akan memberikan rekomendasi terhadap pengguna dengan berdasarkan item atau deskripsi produk yang sudah tersedia. Cara kerja content-based filtering adalah dengan mencari persamaan berdasarkan atribut yang ada pada suatu item produk. Interaksi pengguna terhadap suatu produk akan direkam dan dicari kemiripan tertentu yang akan direkomendasikan kepada pengguna. Terdapat juga platform-platform online yang telah menggunakan teknik content-based filtering untuk meningkatkan minat terhadap produknya, seperti platform streaming lagu spotify, mereka akan merekomendasikan lagu yang sering pengguna dengar dengan melihat informasi ada dalam lagu tersebut seperti penyanyi dan lirik lagu yang memilih deskripsi yang mirip. Terdapat juga platform e-commerce seperti alibaba.com yang menggunakan teknik ini, yaitu dengan cara merekomendasikan produk yang sering pengguna beli dengan melihat kemiripan dari deskripsi produk atau item yang ada dengan produk tersebut.

Salah satu teknik dalam content-based filtering adalah dengan mencari kemiripan data-data multimedia, seperti image, audio atau video. Hal ini menjadi tantangan bagi beberapa peneliti untuk melakukan ekstraksi pada berbagai jenis data tersebut. Salah satu proses ekstraksi gambar adalah menggunakan algoritma CNN (Convolutional Neural Network). Algoritma CNN adalah salah satu cabang algoritma yang ada pada deep learning. Algoritma ini digunakan untuk membaca atau mendeteksi berupa gambar atau data citra digital yang nantinya akan diubah menjadi struktur yang berupa matrix array. Selanjutnya kemiripan antar produk dapat dihitung menggunakan cosine similarity. Cosine similarity adalah salah satu metode yang

digunakan untuk mengukur kemiripan dari sebuah dokumen atau teks. Tujuan utama dalam penelitian ini adalah untuk menghitung matrik kesamaan antara setiap pasangan gambar.

Penelitian ini menggunakan dataset yang berasal dari web kaggle dengan jumlah data 1057 berupa image sepatu futsal. Kemudian data tersebut akan diekstraksi dengan menggunakan metode CNN untuk dijadikan matriks array dan setelah itu akan dicari kemiripan seperti fitur warna, pola tekstur dan juga bentuk produk menggunakan metode cosine similarity. Sehingga nanti dapat dijadikan rekomendasi produk terhadap pengguna yang berbasis image.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka permasalahan yang dapat dirumuskan adalah

- 1) Bagaimana melakukan ekstraksi image untuk sistem rekomendasi berbasis konten
- 2) Bagaimana melakukan evaluasi pada hasil dari sistem rekomendasi

1.3 Batasan Masalah

Beberapa Batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut :

1. Metode yang digunakan untuk rekomendasi adalah content based filtering.
2. Algoritma yang digunakan adalah VGG16 dan Cosine similarity

3. Dataset yang digunakan menggunakan data gambar sepatu futsal yang ada pada sumber terbuka (*Kaggle.com*)
4. Menggunakan metode NDCG (Normalized Discounted Cumulative Gain) untuk proses evaluasinya.

1.4 Maksud dan Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengetahui hasil dari sistem rekomendasi berbasis konten yang diberikan kemudian menghitung nilai evaluasi menggunakan metode NDCG.

1.5 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat bermanfaat untuk memberikan pengetahuan terhadap hasil yang diperoleh dari analisis ekstraksi image pada sistem rekomendasi dengan teknik content-based filtering.

1.6 Metode Penelitian

Berikut adalah tahapan metode penelitian yang akan dilakukan pada skripsi ini

1.6.1 Metode Pengumpulan Data

Data yang akan digunakan dalam penelitian adalah data gambar sepatu futsal yang diambil dari website kaggle <https://www.kaggle.com/ms1437/futsal-shoes>. Data ini berisi 1057 gambar sepatu yang terdiri dari beberapa merk, warna dan bentuknya.

1.6.2 Metode Eksperimen

Dalam tahap ini dilakukan tahap eksperimen terhadap data yang sudah dikumpulkan yang pertama yaitu dengan menginputkan data berupa gambar sepatu futsal kemudian data gambar tersebut diekstrak menggunakan VGG16 yang menghasilkan matriks array. Kemudian dihitung persamaannya dengan menggunakan cosine similarity. Didapatkanlah item yang direkomendasikan oleh sistem. Yang selanjutnya hasil sistem rekomendasinya dievaluasi dengan menggunakan metode NDCG (Normalized Discounted Cumulative Gain).

1.6.3 Metode Pengujian

Pada tahap dilakukan tahap evaluasi terhadap hasil relevansi sistem rekomendasi yang telah diberikan dengan menggunakan metode *NDCG (Normalized Discounted Cumulative Gain)*

1.7 Sistematika Penulisan

Dalam mempermudah menyusun laporan dan mengetahui pembahasan yang ada didalamnya secara menyeluruh, diperlukan sistematika yang merupakan kerangka penulisan skripsi. Berikut adalah sistematika penulisannya laporan ini :

BAB I PENDAHULUAN

Di Dalam bab ini berisi latar belakang, rumusan masalah, batasan masalah, maksud dan tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

BAB II LANDASAN TEORI

Dalam bab ini menjelaskan tentang berbagai teori penunjang dari berbagai referensi seperti buku, jurnal dan laporan penelitian sebelumnya dan berisi dasar teori yang menjelaskan *sistem rekomendasi, content based filtering, algoritma CNN (Convolutional Neural Network) dan VGG16*

BAB III METODE PENELITIAN

Bab ini menjelaskan tentang desain alur penelitian, kemudian metode yang digunakan dan implementasi teknik *content based filtering* pada sistem rekomendasi.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan tentang penjabaran dan pembahasan hasil dari penelitian yang telah dilakukan mengenai implementasi teknik *content based filtering* pada sistem rekomendasi.

BAB V PENUTUP

Bab ini berisi kesimpulan dan saran yang dapat peneliti rangkum selama proses penelitian.

DAFTAR PUSTAKA

Pada bagian ini memuat dari berbagai referensi seperti buku, jurnal atau literatur lain yang menjadi acuan dalam penelitian ini.