

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

*Recommender System* (RS) merupakan komponen perangkat lunak yang menentukan *item* yang sebaiknya ditunjukkan kepada pengguna tertentu berdasarkan statistik dan model machine learning [1]. *Item* merupakan sebutan umum terhadap sesuatu yang ditunjukkan kepada pengguna pada *platform* atau konteks tertentu. *Netflix* menunjukkan film kepada pengguna, sedangkan *Amazon* menunjukkan produk [2]. RS sudah umum digunakan pada berbagai domain di internet, mulai dari *e-commerce*, sosial media, hingga situs berita. Keberadaan RS memudahkan pengguna *platform* untuk menemukan informasi yang relevan secara cepat dalam *volume* informasi yang terus berkembang.

Tujuan utama dari rekomendasi *item* yang ditunjukkan kepada pengguna adalah untuk berisikan pemahaman akurat mengenai kebutuhan pengguna tersebut yang dapat diukur menggunakan berbagai metrik akurasi. Walaupun akurasi digunakan sebagai metrik utama, seiring waktu dibutuhkan juga metrik lainnya untuk mendapatkan gambaran lebih baik mengenai manfaat suatu rekomendasi *item*. Salah satu metrik lainnya adalah keragaman. Keragaman pada umumnya didefinisikan sebagai variasi yang terdapat didalam rekomendasi.

Selain menunjukkan rekomendasi *item* sesuai dengan kebutuhan atau preferensi pengguna, penambahan *item* yang bervariasi dinilai dapat bermanfaat bagi suatu RS karena kebutuhan atau preferensi pengguna dapat berubah seiring waktu [1]. Pengguna dapat merasa tidak puas apabila sudah mengetahui rekomendasi *item* dan merasa tidak mendapat manfaat dari rekomendasi tersebut. Pemilik bisnis pada *platform* juga dapat merasakan kerugian apabila distribusi *item* kurang adil dalam rekomendasi *item* kepada pengguna *platform* [3].

Penelitian ini akan membangun *Hybrid* RS menggunakan algoritma *Alternating Least Squares* (ALS) dan *Categorical Boosting* (CatBoost) untuk menghasilkan rekomendasi *item* kepada pengguna yang akan diukur menggunakan metrik akurasi *Root Mean Squared Error* (RMSE) serta metrik keragaman *Cosine Similarity*. Hasil dari pengukuran akan dibandingkan dengan hasil tersendiri dari algoritma yang membangun *Hybrid* RS. Penelitian dilakukan untuk meningkatkan hasil dari metrik akurasi dan mengukur perubahan yang terjadi pada metrik keragaman.

MyAnimeList dipilih sebagai *platform* studi kasus dengan dasar perkembangan stabil yang dialami oleh industri *anime* pada beberapa tahun terakhir yang juga dapat dilihat dari peningkatan penerapan gaya *anime* pada iklan di Indonesia [4]. Selain peningkatan industri, *anime* juga memiliki kemiripan dengan *medium film* yang sering digunakan pada penelitian RS, contohnya adalah *MovieLens* [5]. Walaupun memiliki *medium* yang serupa, tetapi belum terdapat banyak penelitian yang dilakukan menggunakan *dataset anime* pada konteks RS [6]

## 1.2 Rumusan Masalah

Rumusan masalah pada penelitian adalah sebagai berikut:

- a. Bagaimana membangun Hybrid RS menggunakan algoritma ALS dan CatBoost?
- b. Bagaimana meningkatkan nilai dari metrik akurasi pada *Hybrid* RS ketika dibandingkan dengan algoritma tersendiri?
- c. Bagaimana mengukur akurasi dari *Hybrid* RS menggunakan metrik RMSE?
- d. Bagaimana mengukur keragaman dari *Hybrid* RS menggunakan metrik *Cosine Similarity*?

## 1.3 Batasan Masalah

Batasan masalah pada penelitian adalah sebagai berikut:

- a. *Hybrid* RS akan menggunakan *input dataset* publik *Kaggle* dengan judul *Anime Recommendation Database 2020* dari pengguna Hernan Valdivieso
- b. *Hybrid* RS akan memberikan *output* berupa rekomendasi *anime* dalam bentuk daftar kepada pengguna acak
- c. *Hybrid* RS akan dibangun pada situs web *Google Colab* menggunakan bahasa *python* serta *library* yang tersedia
- d. *Google Colab* akan dijalankan pada *browser* laptop menggunakan akun dasar dengan pemberian *System RAM* sebesar 12 GB

## 1.4 Tujuan Penelitian

Tujuan yang ingin dicapai dari penelitian ini adalah sebagai berikut:

- a. Melakukan pembangunan *Hybrid* RS menggunakan algoritma ALS dan CatBoost
- b. Melakukan peningkatan nilai dari metrik akurasi pada *Hybrid* RS ketika dibandingkan dengan algoritma tersendiri
- c. Melakukan pengukuran akurasi dari *Hybrid* RS menggunakan metrik RMSE
- d. Melakukan pengukuran keragaman dari *Hybrid* RS menggunakan metrik *Cosine Similarity*

## 1.5 Manfaat Penelitian

Manfaat yang ingin didapat dari penerapan penelitian adalah sebagai berikut:

- a. Memberikan *Hybrid* RS yang dapat diterapkan pada *platform* menggunakan *input MyAnimeList*
- b. Memberikan rekomendasi *anime* yang memiliki nilai metrik akurasi RMSE yang lebih baik ketika dibandingkan dengan algoritma tersendiri
- c. Memberikan rekomendasi *anime* yang memiliki nilai metrik keragaman *Cosine Similarity* yang lebih baik ketika dibandingkan dengan algoritma tersendiri