

Hybrid Recommendation System untuk Peminjaman Buku Perpustakaan dengan Collaborative dan Content-Based Filtering

Adrianus Aditya Widjaja, Henry Novianus Palit
Program Studi Informatika, Fakultas Teknologi Industri, Universitas Kristen Petra
Jl. Siwalankerto 121 – 131 Surabaya 60236
Telp. (031) – 2983455, Fax. (031) - 8417658
E – mail: adrianusaditya18@gmail.com, hnpalit@petra.ac.id

ABSTRAK

Seiring dengan perkembangan teknologi yang semakin pesat, perkembangan data digital yang ada di internet juga turut bertambah. Oleh karena itu, sistem rekomendasi ditemukan dan dapat digunakan dalam berbagai aspek kehidupan manusia. Contohnya, sebuah sistem rekomendasi buku yang dapat membantu seseorang untuk memilih buku yang akan dibaca berikutnya sesuai dengan preferensinya. Dengan adanya sistem rekomendasi dapat mengurangi waktu yang dibutuhkan untuk memilih buku karena pilihan buku yang sangat banyak.

Penelitian ini menggunakan Hybrid Recommendation System yang menggabungkan metode Collaborative Filtering dan Content-based Filtering. Penggabungan kedua metode dilakukan dengan tujuan untuk memperoleh hasil rekomendasi yang lebih baik. Untuk mengukur seberapa baik hasil rekomendasi yang dihasilkan akan digunakan metode Mean Reciprocal Rank dan Mean Average Precision. Hasil dari sistem menunjukkan bahwa metode Weighted Hybrid menghasilkan nilai MRR dan MAP yang lebih baik dibanding kedua metode lainnya dengan nilai sebesar 0.2113 dan 0.0988.

Kata Kunci: buku, sistem rekomendasi, collaborative filtering, content-based filtering, sistem rekomendasi hybrid, mean reciprocal rank, mean average precision.

ABSTRACT

The development of technology is growing rapidly and followed with the growth of digital data in the internet. Therefore, recommendation system was invented and can be used in many aspects of human life. For example, a book recommendation system can be used to help user to choose a book to be read which is suitable with their preferences. Using a recommendation system can help to reduce the required time to choose a book because of the massive choices of books.

This research using hybrid recommendation system which combined collaborative filtering and content-based filtering method. The purpose of this study was to achieve a better recommendation outcome. To measure how well the result of the recommendation, mean reciprocal rank and mean average precision was used. The results showed that weighted hybrid yields a better score than the other two methods. The score was 0.2113 and 0.0988 respectively.

Keywords: book, recommendation system, collaborative filtering, content-based filtering, hybrid recommendation system, mean reciprocal rank, mean average precision.

1. PENDAHULUAN

Pesatnya perkembangan teknologi menghasilkan peningkatan jumlah data digital yang terdapat pada internet sehingga menyebabkan peningkatan jumlah informasi yang membuat para pengguna internet kesulitan untuk menentukan informasi apa yang akan digunakan tanpa menghabiskan waktu yang lama [4]. Sistem rekomendasi diperkenalkan pertama kali pada pertengahan tahun 90-an yang merupakan gabungan antara informasi terhadap produk dan algoritma untuk menghasilkan rekomendasi sesuai dengan keinginan pengguna [4, 11]. Sistem rekomendasi dapat memberikan rekomendasi dalam hal – hal rutin seperti, buku, film, dan artikel sampai hal - hal penting lainnya yang membutuhkan berbagai pertimbangan untuk memilihnya seperti, jasa keuangan, mobil, kamera digital, dan apartemen. Terdapat empat metode pada *recommendation system*, yaitu *Collaborative Filtering* (CF), *Content-based Filtering* (CBF), *Demographic Filtering* (DF), dan *Knowledge-based Filtering* (KBF) [4, 6, 12]. Dalam perkembangannya, terdapat sistem rekomendasi yang menggabungkan beberapa metode rekomendasi yang ada, yaitu *hybrid recommendation system*. Tujuannya adalah untuk mengambil kelebihan – kelebihan dari setiap metode yang digunakan dan menyelesaikan permasalahan yang biasanya muncul pada sistem rekomendasi, seperti *scalability*, *sparsity*, dan *cold start problems*.

Penelitian menggunakan metode *hybrid recommendation system* sudah banyak dilakukan. Salah satunya penelitian oleh [3]. Penelitian ini menggabungkan lima metode sistem rekomendasi dan *output*-nya digunakan pada *optimistic exponential ordered weighted averaging* (OWA). Hasil dari metode tersebut lebih baik dibandingkan dengan performa lima sistem rekomendasi yang digunakan lainnya [3]. Peneliti lain yang juga melakukan penelitian pada metode *hybrid recommendation system* adalah [5]. Pada penelitian ini diajukan metode *hybrid* baru yang dinamai HYRED. Hasil dari metode ini juga lebih baik dengan dataset yang sudah disederhanakan dibanding dengan metode lain yang menggunakan dataset utuh [5]. Penelitian lain yang menggunakan metode *hybrid recommendation system* adalah penelitian yang dilakukan oleh [7]. Pada penelitian ini, metode yang digunakan adalah CF dan CBF dengan *Bayesian networks model*. Hasil dari penelitian ini juga menunjukkan hasil yang lebih baik dan dapat diterapkan pada data *real world* [7].

Perpustakaan Universitas Kristen Petra telah memiliki sistem informasi mengenai informasi koleksi buku yang dimiliki. Terdapat dua jenis koleksi buku pada Perpustakaan Universitas Kristen Petra, yaitu *Physical Collection* dalam bentuk fisik dan *Digital Collection* dalam bentuk *digital*. Terdapat lebih dari 160.000 judul koleksi yang dimiliki oleh Perpustakaan Universitas Kristen Petra per September 2014 dan terus mengalami penambahan [10]. Hal

ini akan membuat pengguna perpustakaan mengalami kesulitan dalam menentukan buku apa yang akan dibaca selanjutnya dan buku apa yang sesuai dengan buku yang dibaca, serta membutuhkan waktu yang cukup lama dalam menentukan pilihan. Selain itu, hal ini juga dapat menyebabkan buku – buku yang kurang diminati oleh pengguna menjadi semakin tidak diketahui dan menyebabkan kerugian terhadap perpustakaan.

Untuk mengatasi masalah diatas maka dibutuhkan sebuah sistem rekomendasi yang dapat membantu pengguna perpustakaan untuk menemukan buku yang dibutuhkan atau berhubungan untuk dibaca selanjutnya. Selain itu, dengan adanya sistem rekomendasi ini diharapkan dapat meningkatkan jumlah peminjaman buku dan lebih banyak lagi buku – buku yang sebelumnya masih kurang terekspos menjadi dibaca atau dipinjam oleh banyak pengguna lainnya. Penelitian ini menggabungkan metode CF dan CBF karena rekomendasi yang diberikan disesuaikan dengan kemiripan selera dan preferensi buku sesuai dengan pengguna lain yang memiliki kemiripan serupa.

Penelitian ini tidak menggunakan dataset perpustakaan Universitas Kristen Petra secara langsung dikarenakan pada dataset perpustakaan tidak terdapat *rating* terhadap buku yang dibaca sehingga metode yang digunakan tidak dapat diterapkan. Dataset goodreads dari kaggle dipilih karena memiliki kesamaan dengan data perpustakaan Universitas Kristen Petra pada tabel *books* dan *tags* (judul buku, penulis, ISBN, *genre*, *tag*, dan lain - lain) sehingga dataset dapat merepresentasikan dataset perpustakaan. Kelemahan dataset ini, yaitu pengguna pada dataset ini berasal dari populasi yang luas karena berasal dari *web* goodreads. Sedangkan, populasi dataset pada perpustakaan Universitas Kristen Petra mencakup pengguna di lingkungan kampus. Hal ini akan mengakibatkan perbedaan terhadap preferensi pengguna dan tingkat kepuasan pengguna. Model sistem rekomendasi ini dapat diterapkan pada perpustakaan Universitas Kristen Petra jika dataset perpustakaan Universitas Kristen Petra menambahkan penggunaan *rating* yang diberikan oleh pengguna terhadap buku. Penelitian ini diharapkan dapat menjadi langkah awal terhadap pembaharuan sistem informasi dan rekomendasi buku di Perpustakaan Universitas Kristen Petra. Selain itu juga bisa memberikan sistem rekomendasi yang lebih cepat dan sesuai dengan preferensi pengguna.

2. LANDASAN TEORI

2.1 Perpustakaan

Buku merupakan salah satu aspek yang membutuhkan sistem rekomendasi untuk memberikan rekomendasi dari banyaknya pilihan dan informasi yang tersedia sehingga sesuai dengan preferensi pengguna. Berdasarkan data Google pada tahun 2010, terdapat 129.864.880 juta judul buku dan setidaknya terdapat 1 juta judul baru yang terbit di UK, China, dan Amerika pada tahun 2013. Keterangan yang ada pada buku biasanya terdiri dari judul buku, sinopsis, penulis, tahun publikasi, publisher, genre buku, dan lain-lain [2]. Buku memiliki beberapa jenis genre, seperti misteri, science, fiksi sastra, fiksi ilmiah, dan lain-lain [13]. Menurut International Federation of Library Associations and Institutions, perpustakaan merupakan kumpulan informasi yang tersusun secara sistematis untuk memudahkan pengguna, berupa tercetak, non-cetak, dan/atau informasi dalam computer [1].

2.2 Sistem Rekomendasi

Sistem rekomendasi merupakan sebuah sistem yang dibuat pada pertengahan tahun 90-an. Sistem ini dikembangkan untuk membantu menyediakan pilihan rekomendasi barang (seperti film, buku, lagu, produk, dan lainnya) atau pelayanan sesuai dengan

kebutuhan pengguna. Pada awalnya, sistem ini ditujukan dalam penelitian yang berhubungan dengan *Human Computer Interaction* (HCI) dan *Information Retrieval* (IR). Terdapat beberapa macam metode dari sistem rekomendasi, yaitu CF, CBF, DF, dan KBF [4, 6, 12].

2.3 Collaborative Filtering

CF merupakan metode yang menggunakan data dari pengguna lain yang memiliki kemiripan selera dengan target pengguna dan memberikan rekomendasi berdasarkan kemiripan tersebut. Hal ini dapat digambarkan dalam kemiripan kesukaan atau rating antara pengguna lain dan target pengguna. Di dalam metode CF terdapat dua macam pendekatan, yaitu *user-based* dan *item-based* [14].

CF memiliki dua metode, yaitu *memory-based* dan *model-based*. Metode *memory-based* merupakan metode yang menghasilkan prediksi *rating* yang baru berdasarkan keseluruhan koleksi items yang telah diberi *rating* oleh pengguna lain. Sedangkan, metode *model-based* merupakan yang menghasilkan prediksi *rating* yang baru dengan menggunakan data dari seluruh *rating* yang telah disederhanakan ukurannya [14].

Salah satu langkah didalam perhitungan algoritma CF adalah menghitung *similarity* antara *items* dan *users* dan memilih kelompok rekomendasi yang paling mendekati dengan pengguna. Salah satu metode perhitungan yang dapat digunakan untuk menghitung *similarity* didalam CF:

2.3.1 Cosine Similarity

Metode *cosine similarity* merupakan metode yang menghitung kemiripan antara dua benda *a* dan *b* yang memiliki dua vektor untuk menghitung *Cosine Vector* (CV) dimana kemiripan diantara dua vektor menandakan jarak diantara keduanya [14].

$$CS(u, v) = \cos(X_u, X_v) = \frac{\sum_{i \in L_{uv}} r_{ui} r_{vi}}{\sqrt{\sum_{i \in L_u} r_{ui}^2} \sqrt{\sum_{j \in L_v} r_{vj}^2}} \quad (1)$$

2.4 Content-based Filtering

Sistem rekomendasi *content-based* merupakan sistem rekomendasi yang menggunakan kelompok *items* yang telah mendapat *rating* oleh pengguna untuk menentukan preferensi dari pengguna tersebut. Preferensi pengguna tersebut dapat dibentuk melalui kata kunci (*keywords*), kategori, maupun atribut lain dari *item* yang digunakan. Setelah itu, akan dilakukan pencarian *item – item* lain yang memiliki kemiripan dengan preferensi pengguna untuk direkomendasikan [9, 11].

2.4.1 Term Frequency Inverse Document Frequency (TF - IDF)

Terdapat metode yang dapat digunakan untuk mengidentifikasi kata kunci yang terdapat pada data CBF dengan menggunakan metode TF – IDF [11].

$$TF(d,t) = \begin{cases} 0 \\ 1 + \log(1 + \log freq(d,t)) \end{cases} \quad (2)$$

TF digunakan untuk menghitung frekuensi sebuah kata pada data.

$$IDF(t) = \log \frac{|d|}{|dt|} \quad (3)$$

IDF digunakan untuk menentukan seberapa penting kata tersebut dan nilainya berkurang jika kata tersebut semakin banyak muncul pada data.

$$TF-IDF(dt) = TF(d,t) * IDF(t) \quad (4)$$

Setelah terbentuk preferensi pengguna akan dicari *item - item* lain yang memiliki kemiripan dengan preferensi pengguna dengan metode *similarity*. Lalu, *item - item* yang nilai kemiripannya tinggi akan direkomendasikan.

2.5 Hybrid Recommendation System

Sistem rekomendasi *hybrid* adalah sistem rekomendasi yang menggabungkan beberapa metode sistem rekomendasi yang ada menjadi satu. Tujuan dari penggabungan ini adalah untuk mendapatkan kelebihan dan saling menutupi kekurangan dari masing – masing metode sistem rekomendasi yang dipilih.

Terdapat tujuh macam metode dalam *hybrid recommendation*, yaitu *weighted*, *switching*, *mixed*, *feature combination*, *cascade*, *feature augmentation*, dan *meta-level*. *Weighted hybrid recommendation* merupakan penggabungan nilai dari metode rekomendasi yang digunakan. Setelah itu diurutkan dari yang tertinggi ke terendah. *Switching hybrid recommendation* akan melakukan pergantian metode rekomendasi yang digunakan berdasarkan keadaan sekarang dan dilakukan oleh sistem. *Mixed hybrid recommendation* akan menampilkan hasil rekomendasi dari beberapa sistem rekomendasi secara bersamaan. *Feature combination* akan menggunakan informasi dari sistem rekomendasi sebagai data tambahan bagi sistem rekomendasi lainnya. *Cascade hybrid recommendation* tidak akan menggunakan *item - item* yang sudah baik hasilnya atau tidak mungkin direkomendasikan pada sistem rekomendasi pertama pada sistem rekomendasi kedua. *Feature augmentation* menggunakan *output* dari sistem rekomendasi pertama menjadi *input* untuk sistem rekomendasi berikutnya. *Meta-level* menggunakan keseluruhan model dari sistem rekomendasi pertama sebagai *input* dari sistem rekomendasi berikutnya [9].

2.6 Data Sampling

Data sampling adalah metode yang digunakan untuk mengambil sampel dari populasi karena keterbatasan waktu atau sumber daya untuk menganalisa seluruh populasi. Terdapat dua metode utama pada proses *data sampling* yaitu, *random sampling* dan *non-random sampling*. Pada *random sampling* terdapat metode *simple random sampling*, *stratified random sampling*, *cluster sampling*, *systematic random sampling*, dan *multi stage sampling*. Sedangkan pada *non-random sampling* terdapat metode *quota sampling*, *snowball sampling*, *judgement sampling*, dan *convenience sampling* [15].

2.7 Mean Reciprocal Rank (MRR)

Mean Reciprocal Rank merupakan sebuah metode untuk mengevaluasi performa dari sistem rekomendasi yang dibuat. MRR bekerja dengan cara menilai setiap *item* yang relevan berdasarkan posisinya pada rekomendasi dan menjumlahkannya.

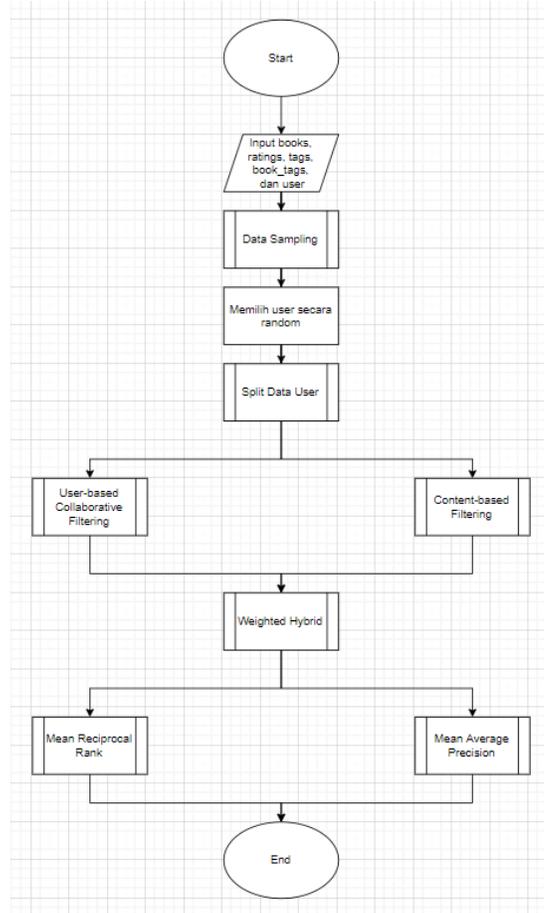
$$MRR = \frac{1}{|Q|} \sum_{i=1}^{|Q|} \frac{1}{rank_i} \tag{5}$$

2.8 Mean Average Precision (MAP)

Mean Average Precision merupakan sebuah metode perhitungan yang digunakan untuk mengevaluasi performa dari sebuah sistem rekomendasi. MAP diperoleh dari rata – rata seluruh nilai *average precision* yang dihasilkan percobaan.

$$MAP = \frac{\sum_{q=1}^Q \frac{sum(R_q \cap T_q)}{sum(R_q)}}{Q} \tag{6}$$

3. DESAIN SISTEM



Gambar 1. Flowchart Hybrid Recommendation System

Proses flowchart pada Gambar 1 dimulai dengan *input* tabel *books*, *ratings*, *tags*, dan *book_tags*. Tabel – tabel tersebut akan dimasukkan kedalam data *frame* dan dilakukan proses *data sampling*. Dari total *user* diambil sebanyak 10000 *user* dengan metode *stratified random sampling* berdasarkan jumlah *rating* yang diberikan oleh *user*. Setelah itu dilakukan pemilihan *user* secara *random* untuk pengujian dan data *rating user* di-split menjadi data *training* dan data *testing*. Selanjutnya, dilakukan proses *Collaborative Filtering* dan *Content-based Filtering*. Kedua proses tersebut akan dilakukan secara terpisah dan hasilnya akan digabungkan menjadi sebuah sistem rekomendasi *Hybrid*. Setelah itu, buku yang akan direkomendasikan berasal dari sistem rekomendasi *Hybrid* tersebut.

Pada *Collaborative Filtering*, proses pertama membentuk *user rating matrix* berdasarkan *dataframe* pada tabel *ratings*. Data pada *dataframe* akan dirubah menjadi *pivot table* dengan *book_id* dan *user_id* sebagai baris dan kolom beserta *rating* sebagai *value*-nya. Buku yang telah diberi *rating* oleh *user* *value*-nya akan berisi nilai *rating*-nya, sedangkan buku yang belum diberi *rating* oleh *user* *value*-nya akan berisi NaN dan digantikan dengan nilai rata-rata *rating* masing-masing *user* [8].

Setelah *pivot table* terbentuk langkah selanjutnya mencari *similarity* antara *user* yang akan diberi rekomendasi buku dengan

seluruh *user* lainnya menggunakan *cosine similarity*. Berikutnya akan dipilih *nearest neighbor* dengan jumlah $K = 100$ *user* serta 20 dan 500 *user* sebagai pembanding. Hal ini dilakukan untuk mencari jumlah *nearest neighbor* yang sesuai digunakan karena setiap *dataset* yang digunakan memiliki jumlah *nearest neighbor* yang berbeda-beda untuk menghasilkan rekomendasi yang terbaik.

Setelah ditentukan *nearest neighbor* yang akan digunakan untuk memberikan rekomendasi buku akan dihitung prediksi *rating* untuk buku yang belum diberi *rating* berdasarkan pada *rating* buku dari *nearest neighbor* yang digunakan dengan diberi bobot berdasarkan *similarity*-nya. Hasil tersebut akan digunakan dari nilai tertinggi sebagai rekomendasi yang diberikan kepada *user*.

Pada *Content-based Filtering*, preferensi *user* dibentuk berdasarkan *tag* buku yang dibaca oleh *user* tersebut. Untuk membentuknya akan digunakan *Term Frequency – Inverse Document Frequency* (TF – IDF). TF akan digunakan untuk menghitung seberapa banyak sebuah *tag* muncul. IDF akan digunakan untuk mengurangi nilai *term* yang tidak terlalu penting seperti kata penghubung. Setelah preferensi *user* telah terbentuk, langkah berikutnya adalah mencari *similarity* setiap buku yang belum dibaca oleh *user* tersebut terhadap preferensi *user* menggunakan *cosine similarity*.

Proses pada metode *Hybrid* diawali dengan melakukan normalisasi pada hasil prediksi *rating* yang dihasilkan metode *Collaborative Filtering* dari *range* 1-5 menjadi 0-1. Proses normalisasi dilakukan untuk menyamakan *range* antara hasil *Collaborative Filtering* dan *Content-based Filtering* yang berupa nilai *similarity* 0-1. Setelah itu, hasil normalisasi dan *Content-based Filtering* digabungkan untuk masing-masing buku. Setelah itu, dilakukan proses *performance measurement* dengan *evaluation metric Mean Reciprocal Rank* dan *Mean Average Precision*.

4. PENGUJIAN SISTEM

4.1 Dataset

Sistem ini menggunakan dua dataset, yaitu dataset goodreads dari kaggle dan data perpustakaan Universitas Kristen Petra. Pada data perpustakaan Universitas Kristen Petra terdapat 20.642 *master* buku dan 24.457 buku yang terbit antara tahun 2002-2012 serta 24.534 data peminjaman sepanjang tahun 2014.

4.2 Proses Pengujian

Pengujian menggunakan data *sampling* sebanyak 10000 *user* yang dilakukan dengan metode *stratified random sampling* berdasarkan jumlah *rating* yang diberikan dan diterapkan pada tiga metode, yaitu *Hybrid* serta *User-based Collaborative Filtering* dan *Content-based Filtering* sebagai metode pembanding. Pada metode *User-based Collaborative Filtering* digunakan *nearest neighbor* sebanyak 100 *user* serta 20 dan 500 *user* sebagai pembanding. Untuk proses pengujian data *rating user* yang diuji dibagi menjadi data *training* dan *testing*. Hasil rekomendasi yang diberikan berjumlah 50 buku dan ditampilkan dalam bentuk 10 buku per halaman. Lalu, hasil dari masing – masing metode akan dihitung seberapa baik hasil rekomendasi yang diberikan dengan *Mean Reciprocal Rank* dan *Mean Average Precision*. *Mean Reciprocal Rank* digunakan untuk melihat seberapa tinggi posisi atau *ranking* dari rekomendasi buku yang diberikan dan terdapat pada data *testing*. Sedangkan, *Mean Average Precision* digunakan untuk mengetahui berapa banyak buku dari hasil rekomendasi yang berada pada data *testing*.

4.3 Hasil Mean Reciprocal Rank dan Mean Average Precision

Hasil pengujian merupakan hasil rata-rata dari tiga set pengujian dan setiap set berisi 110 *user*. Proses pengujian dilakukan sebanyak tiga kali agar hasil yang didapatkan lebih kredibel dibandingkan dengan melakukan pengujian sebanyak satu kali saja.

Untuk $K = 20$, nilai MRR yang dihasilkan = 0.0417 dan nilai MAP = 0.0182 untuk metode *Collaborative*. Sedangkan untuk metode *Hybrid*, nilai yang dihasilkan sebesar 0.2087 untuk MRR dan 0.0988 untuk MAP. Untuk metode *Content-based* nilai MRR dan MAP yang dihasilkan sebesar 0.1786 dan 0.0897. metode *Content-based* hanya memiliki satu hasil karena tidak dipengaruhi oleh variasi jumlah *nearest neighbor* yang dilakukan pada pengujian.

Untuk *nearest neighbor* sebanyak 100 *user*, metode *Collaborative* menghasilkan nilai MRR dan MAP sebesar 0.1623 dan 0.0764. Metode *Hybrid* menghasilkan nilai sebesar 0.2113 dan 0.0988 untuk MRR dan MAP secara berurutan.

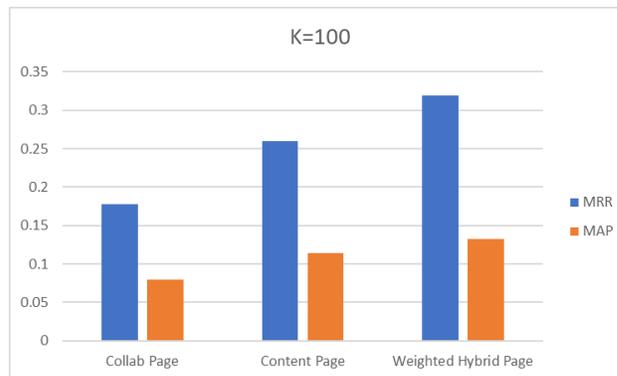
Metode *Collaborative* dengan *nearest neighbor* 500 *user* menghasilkan nilai MRR dan MAP sebesar 0.1901 dan 0.0988. Untuk metode *Hybrid* menghasilkan nilai sebesar 0.221 dan 0.1061 untuk nilai MRR dan MAP secara berurutan. Hasil keseluruhan untuk MRR dan MAP dapat dilihat pada Tabel 1.

Tabel 1. Hasil Pengujian MRR dan MAP

Metode	Testing Data (3 set, n = 110)					
	K = 20		K = 100		K = 500	
	MRR	MAP	MRR	MAP	MRR	MAP
<i>Collaborative</i>	0.0417	0.0182	0.1623	0.0764	0.1901	0.0988
<i>Content-based</i>	0.1786	0.0897	0.1786	0.0897	0.1786	0.0897
<i>Hybrid</i>	0.2087	0.0988	0.2113	0.0988	0.221	0.1061

Gambar 2 menunjukkan hasil metode yang diusulkan yaitu, metode *Hybrid* dengan *nearest neighbor* 100 *user* merupakan hasil terbaik kedua dengan selisih 0.0097 dan 0.0073 lebih rendah untuk nilai MRR dan MAP secara berurutan dibandingkan dengan metode *Hybrid* dengan *nearest neighbor* 500 *user*. Walaupun memiliki hasil terbaik, metode *Hybrid* dengan *nearest neighbor* 500 *user* terkendala dengan waktu *running* sistem rekomendasi yang lebih lama.

Pada Tabel 2 menunjukkan hasil perbandingan waktu *running* sistem rekomendasi menggunakan ketiga metode dengan variasi *nearest neighbor* yang ditentukan. Walaupun metode *Hybrid* dengan *nearest neighbor* 500 *user* menghasilkan nilai MRR dan MAP yang terbaik membutuhkan waktu *running* yang paling lama



Gambar 2. Grafik Mean Reciprocal Rank dan Mean Average Precision menggunakan $K=100$

dibandingkan dengan metode yang lain. Sedangkan metode *Hybrid* dengan *nearest neighbor* 100 *user* yang menghasilkan hasil terbaik kedua dengan selisih nilai yang kecil membutuhkan waktu *running* 48% lebih cepat dibandingkan dengan metode *Hybrid* dengan *nearest neighbor* 500 *user*. Sehingga, dengan memperhitungkan hasil nilai MRR dan MAP serta waktu *running* sistem rekomendasi dapat disimpulkan bahwa metode yang diusulkan lebih baik.

Tabel 2. Perbandingan Waktu Running Sistem Rekomendasi

Metode	Jumlah <i>nearest neighbor</i> (<i>user</i>)	Rata-rata waktu (s)
<i>Collaborative</i>	20	21.717
	100	25.971
	500	53.124
<i>Content-based</i>	-	3.579
<i>Hybrid</i>	20	25.846
	100	29.441
	500	56.421

Gambar 3 menunjukkan hasil MRR dan MAP dari 32 *user* pada *dataset* perpustakaan dengan metode *Content-based Filtering*. Perhitungan terhadap data perpustakaan hanya dilakukan dengan menggunakan metode *Content-based* karena tidak terdapat data *rating*. Nilai MRR dan MAP menunjukkan nilai sebesar 0.0547 dan 0.0313 secara beurutuan.



Gambar 3. Grafik Mean Reciprocal Rank dan Mean Average Precision pada Data Perpustakaan

Berdasarkan hasil diskusi dengan Ibu Dian Wulandari selaku Kepala Perpustakaan Universitas Kristen Petra dan Bapak Billy Setyadi selaku Kepala Bidang Pendayagunaan Informasi Perpustakaan Universitas Kristen Petra, dapat disimpulkan bahwa dibutuhkan sebuah sistem rekomendasi pada Perpustakaan Universitas Kristen Petra untuk mempromosikan koleksi buku yang dimiliki oleh Perpustakaan. Dengan adanya sistem rekomendasi diharapkan dapat membantu pengguna perpustakaan untuk mengeksplor buku-buku lain yang belum pernah dipinjam. Untuk pengadaan data *rating* memiliki tantangan tersendiri karena jumlahnya yang sedikit. Tetapi, dengan adanya data *rating* diharapkan dapat meningkatkan jumlah peminjaman buku di perpustakaan.

5. KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari hasil pengujian sistem yang telah dilakukan, dapat diambil beberapa kesimpulan, yaitu:

- Berdasarkan nilai MRR dan MAP pada metode *Hybrid* dengan $K = 100$ sebesar 0.2113 dan 0.0988 menunjukkan

bahwa hasil rekomendasi yang diberikan masih memiliki nilai yang jauh dari 1 sehingga masih bisa dioptimalkan untuk memberikan hasil rekomendasi yang sesuai dengan preferensi pengguna.

- Metode *Hybrid* pada $K = 100$ menghasilkan hasil MRR dan MAP lebih baik dengan nilai 0.2113 dan 0.0988 dibandingkan *Collaborative filtering* dengan nilai 0.1623 dan 0.0764 dan *Content-based filtering* dengan nilai 0.1786 dan 0.0897 yang digunakan sebagai metode pembandingan. Tetapi, masih lebih rendah dibandingkan dengan metode *Hybrid* dengan $K = 500$ dengan selisih hasil MRR dan MAP sebesar 0.0097 dan 0.0073.
- Metode *Hybrid* pada $K = 100$ membutuhkan waktu rata-rata *running* selama 29.441 detik yang 48% lebih cepat dibandingkan dengan metode *Hybrid* dengan $K = 500$.

5.2 Saran

Berdasarkan hasil dari pengujian sistem, Saran yang dapat diberikan untuk menyempurnakan dan mengembangkan sistem rekomendasi lebih lanjut antara lain:

- Pada metode *Collaborative Filtering* dilakukan clustering untuk mengatasi data *sparsity* dan menghemat waktu *running* sistem rekomendasi.
- Penggunaan jumlah *nearest neighbor* tidak disamakan untuk seluruh *user*.
- Pada metode *Content-based Filtering* menggunakan atribut buku lainnya dalam membentuk preferensi *user* seperti, judul buku dan penulis.
- Mengubah metode perhitungan *Hybrid* untuk mendapatkan hasil rekomendasi yang lebih baik.
- Metode *Hybrid* dapat diimplementasikan pada sistem rekomendasi Perpustakaan Universitas Kristen Petra.

6. DAFTAR PUSTAKA

- Adistia, L. D., Akhriza, T. M., & Jatmiko, S. 2019. Sistem Rekomendasi Buku untuk Perpustakaan Perguruan Tinggi Berbasis Association Rule. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 3(2), 304–312. <https://doi.org/10.29207/resti.v3i2.971>
- Alharthi, H., Inkpen, D., & Szpakowicz, S. 2018. A survey of book recommender systems. *Journal of Intelligent Information Systems*, 51(1), 139–160. <https://doi.org/10.1007/s10844-017-0489-9>
- Basiri, J., Shakery, A., Moshiri, B., & Zi Hayat, M. 2010. Alleviating the cold-start problem of recommender systems using a new hybrid approach. *2010 5th International Symposium on Telecommunications, IST 2010*, 962–967. <https://doi.org/10.1109/ISTEL.2010.5734161>
- Çano, E., & Morisio, M. 2017. Hybrid recommender systems: A systematic literature review. *Intelligent Data Analysis*, 21(6), 1487–1524. <https://doi.org/10.3233/IDA-163209>
- Choi, S. H., Jeong, Y. S., & Jeong, M. K. 2010. A hybrid recommendation method with reduced data for large-scale application. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 40(5), 557–566. <https://doi.org/10.1109/TSMCC.2010.2046036>
- Das, D., Sahoo, L., & Datta, S. 2017. A Survey on Recommendation System. *International Journal of Computer Applications*, 160(7), 6–10. <https://doi.org/10.5120/ijca2017913081>

- [7] De Campos, L. M., Fernández-Luna, J. M., Huete, J. F., & Rueda-Morales, M. A. 2010. Combining content-based and collaborative recommendations: A hybrid approach based on Bayesian networks. *International Journal of Approximate Reasoning*, 51(7), 785–799. <https://doi.org/10.1016/j.ijar.2010.04.001>
- [8] Desrosiers, C., & Karypis, G. 2011. Recommender Systems Handbook. In *Recommender Systems Handbook*. <https://doi.org/10.1007/978-0-387-85820-3>
- [9] Felfernig, A., Jeran, M., Ninaus, G., Reinfrank, F., Reiterer, S., & Stettinger, M. 2014. Basic approaches in recommendation systems. *Recommendation Systems in Software Engineering*, 15–37. https://doi.org/10.1007/978-3-642-45135-5_2
- [10] *Koleksi dan Gedung*. 2014. https://library.petra.ac.id/index.php?r=site/koleksi_gedung
- [11] Manjula, R., & Chilambuchelvan, A. 2016. Content Based Filtering Techniques in Recommendation System using user preferences. *International Journal of Innovations in Engineering and Technology*, 7(4), 151.
- [12] Mathew, P., Kuriakose, B., & Hegde, V. 2016. Book Recommendation System through content based and collaborative filtering method. *Proceedings of 2016 International Conference on Data Mining and Advanced Computing, SAPIENCE 2016*, 47–52. <https://doi.org/10.1109/SAPIENCE.2016.7684166>
- [13] Mooney, R. J., & Roy, L. 2000. Content-based book recommending using learning for text categorization. *Proceedings of the ACM International Conference on Digital Libraries*, 195–204. <https://doi.org/10.1145/336597.336662>
- [14] Nilashi, M., Bagherifard, K., Ibrahim, O., Alizadeh, H., Nojeem, L. A., & Roozegar, N. 2013. Collaborative filtering recommender systems. *Research Journal of Applied Sciences, Engineering and Technology*, 5(16), 4168–4182. <https://doi.org/10.19026/rjaset.5.4644>
- [15] Taherdoost, H. 2018. Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research. *SSRN Electronic Journal*, 5(2), 18–27. <https://doi.org/10.2139/ssrn.3205035>