

**SISTEM REKOMENDASI BUKU BERBASIS USER COLLABORATIVE
FILTERING DENGAN ALGORITMA K-NEAREST NEIGHBOR**

SKRIPSI



disusun oleh

Arfian Dimas Andi Permana

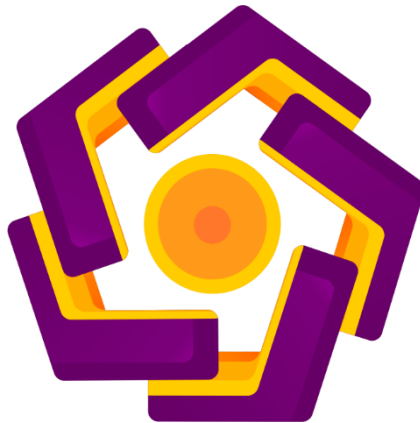
17.11.1018

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

**SISTEM REKOMENDASI BUKU BERBASIS USER COLLABORATIVE
FILTERING DENGAN ALGORITMA K-NEAREST NEIGHBOR**

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi informatika



disusun oleh

Arfian Dimas Andi Permana

17.11.1018

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2021**

PERSETUJUAN

SKRIPSI

**SISTEM REKOMENDASI BUKU BERBASIS USER COLLABORATIVE
FILTERING DENGAN ALGORITMA K-NEAREST NEIGHBOR**

yang dipersiapkan dan disusun oleh

Arfian Dimas Andi Permana
17.11.1018

telah disetujui oleh Dosen Pembimbing Skripsi

pada tanggal 28 Oktober 2020

Dosen Pembimbing,

Arif Dwi Laksito, M.Kom

NIK. 190302150

PENGESAHAN
SKRIPSI
SISTEM REKOMENDASI BUKU BERBASIS USER COLLABORATIVE
FILTERING DENGAN ALGORITMA K-NEAREST NEIGHBOR

yang dipersiapkan dan disusun oleh

Arfian Dimas Andi Permana

17.11.1018

telah dipertahankan di depan Dewan Penguji

pada tanggal 18 November 2021

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Nur'aini, M.Kom

NIK. 190302066

Rakhma Shafrida Kurnia, S.Kom., M.Kom

NIK. 190302355

Arif Dwi Laksito, M.Kom

NIK. 190302150

Skripsi ini telah diterima sebagai salah satu persyaratan

untuk memperoleh gelar Sarjana Komputer

Tanggal 04 Desember 2021

DEKAN FAKULTAS ILMU KOMPUTER

Hanif Al Fatta, S.Kom., M.Kom.

NIK. 190302096

PERNYATAAN

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 26 November 2021



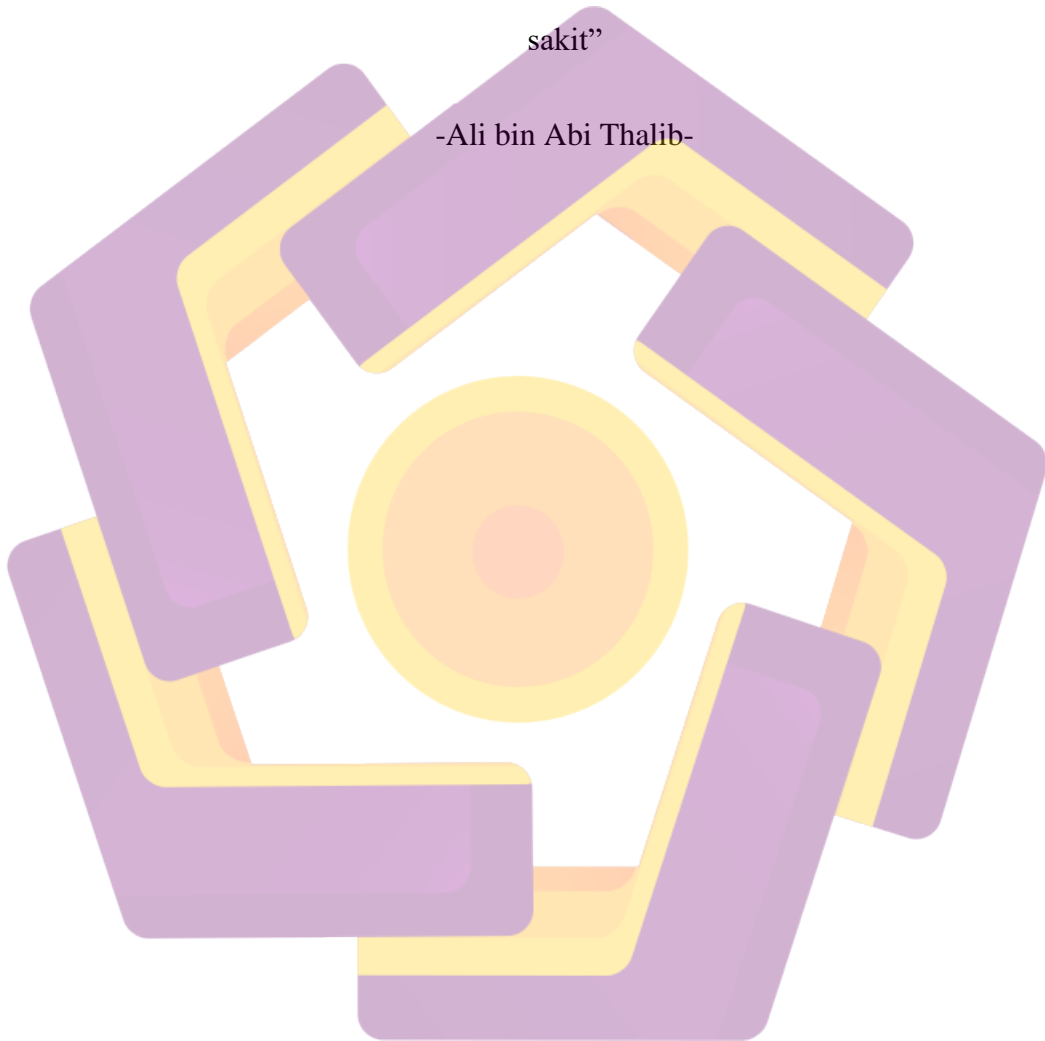
Arfian Dimas Andi Permana

NIM. 17.11.1018

MOTTO

“Yakinlah ada sesuatu yang menantimu selepas banyak kesabaran (yang kau jalani), yang akan membuatmu terpana hingga kau lupa betapa pedihnya rasa sakit”

-Ali bin Abi Thalib-



PERSEMBAHAN

Puji syukur kita panjatkan kehadirat Allah SWT atas berkah dan karunia-Nya skripsi ini dapat terselesaikan dengan baik dan lancar. Dengan ini saya persembahkan skripsi ini untuk semua pihak yang terlibat langsung maupun tidak langsung, yaitu kepada :

1. Kedua orang tua tercinta dan tersayang terimakasih atas segala doa yang dipanjatkan kepada Allah SWT dan dukungan secara moril dan material.
2. Keluarga tercinta dan tersayang (Dinda dan Fadil) terimakasih atas semua doa dan motivasi yang telah diberikan.
3. Aprilia Kurnia Putri terimakasih atas segala bantuan dan dukungan yang telah diberikan.
4. Teman-teman Bracketbrick yaitu Zauvik Rizaldi, Aji Syahroni, Muhammad Aji Putra, dan Fendi Rahman yang telah memberikan dukungan.

Sekali lagi penulis ucapkan terimakasih semoga kebaikan kalian menjadi amal ibadah dan semoga Allah SWT membalas semua kebaikan kalian. Aamiin.

KATA PENGHANTAR

Puji syukur penulis ucapkan kepada Allah SWT, yang mana telah memberikan kesehatan dan karunia-Nya kepada penulis serta kekuatan untuk menyelesaikan skripsi yang berjudul “ **SISTEM REKOMENDASI BUKU BERBASIS USER COLLABORATIVE FILTERING DENGAN ALGORITMA K-NEAREST NEIGHBOR**”. Tidak lupa penulis mengucapkan shalawat dan salam kepada junjungan Nabi Besar Muhammad SAW. Penyelesaian tulisan ini terlepas bantuan dari berbagai pihak yang terkait secara langsung maupun tidak langsung, terutama dan istimewa dipersembahkan kepada kedua orang tua tercinta yang senantiasa memberikan rasa sayang, didikan, serta doa yang selalu di panjatkan pada Allah kepada penulis.

Skripsi ini dapat terselesaikan dengan bantuan berbagai pihak, maka dari itu penulis menyatakan rasa hormat dan terimakasih kepada:

1. Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Bapak Hanif Al Fatta, S.Kom., M.Kom. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Ibu Windha Mega Pradnya D, M.Kom selaku ketua Program Studi Informatika Universitas AMIKOM Yogyakarta.
4. Bapak Arif Dwi Laksito, M.Kom selaku pembimbing yang senantiasa memberikan masukan serta nasihat dalam penulisan skripsi ini.

5. Ibu Nur'aini, M.Kom dan Ibu Rakhma Shafrida Kurnia, S.Kom., M.Kom selaku dosen penguji, terima kasih atas saran dan kritiknya sehingga penelitian ini menjadi lebih baik lagi.

Penulis menyadari masih ada kekurangan dan kelemahan dalam pembuatan skripsi ini. Maka penulis mengharapkan adanya kritik dan saran dari segala pihak agar menambah kesempurnaan dalam skripsi ini.

Yogyakarta, 04 Desember 2021

Arfian Dimas Andi Permana

NIM. 17.11.1018

DAFTAR ISI

PERSETUJUAN	iii
PENGESAHAN	iv
PERNYATAAN	v
MOTTO	vi
PERSEMBAHAN	vii
KATA PENGHANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvii
INTISARI	xviii
ABSTRACT	xix
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	4
1.4 Maksud dan Tujuan Penelitian.....	4
1.5 Manfaat Penelitian	5
1.6 Metode Peneletian.....	5
1.6.1 Metode Studi literatur	5

1.6.2	Metode Data Mining	6
1.6.3	Metode Pengujian.....	7
1.7	Sistematika Penulisan	7
BAB II LANDASAN TEORI		9
2.1	Tinjauan Pustaka	9
2.2	Data Mining	15
2.2.1	Tahap data Mining	15
2.2.2	Pengelompokan Data Mining.....	17
2.3	Sistem Rekomendasi	19
2.4	Content-Based Filtering	19
2.5	Collaborative Filtering	20
2.5.1	User-Based Collaborative Filtering.....	21
2.5.2	Item-Based Collaborative Filtering.....	22
2.6	Hybird Filtering.....	23
2.7	K-Nearest Neighbor	23
2.7.1	Nilai Bobot User Similarity	24
2.8	Nilai Rekomendasi Item.....	25
2.9	Nilai Predeksi Rating Item	25
2.10	Mean Absolute Error (MAE)	26
2.11	Mean Squared Error (MSE)	27

2.12	Robot Mean Squared Error (RMSE).....	27
2.13	Kaggle	28
2.14	Google Collaboratory.....	28
2.15	Diagram Alir	29
BAB III METODE PENELITIAN		30
3.1	Alur Penelitian	30
3.1.1	Pengambilan Data	32
3.1.2	Pemilihan Data.....	32
3.1.3	Pembersihan data	32
3.1.4	Analisa Data.....	33
3.1.5	Pengolahan Data.....	33
3.1.5.1	Perhitungan Manual Pengolahan Data.....	35
3.1.5.1.1	Menentukan User X dan User Y	35
3.1.5.1.2	Proses K-Nearest Neighbor.....	39
3.1.5.1.2.1	Perhitungan Euclidean Distance.....	39
3.1.5.1.2.2	Menghitung Bobot User Y	41
3.1.5.1.2.3	Menentukan Nilai K dan Mengambil Data User Berdasarkan Nilai K.....	43
3.1.5.1.3	Nilai Rekomendasi Buku.....	44
3.1.5.1.4	Nilai Predeksi Rating Buku.....	46

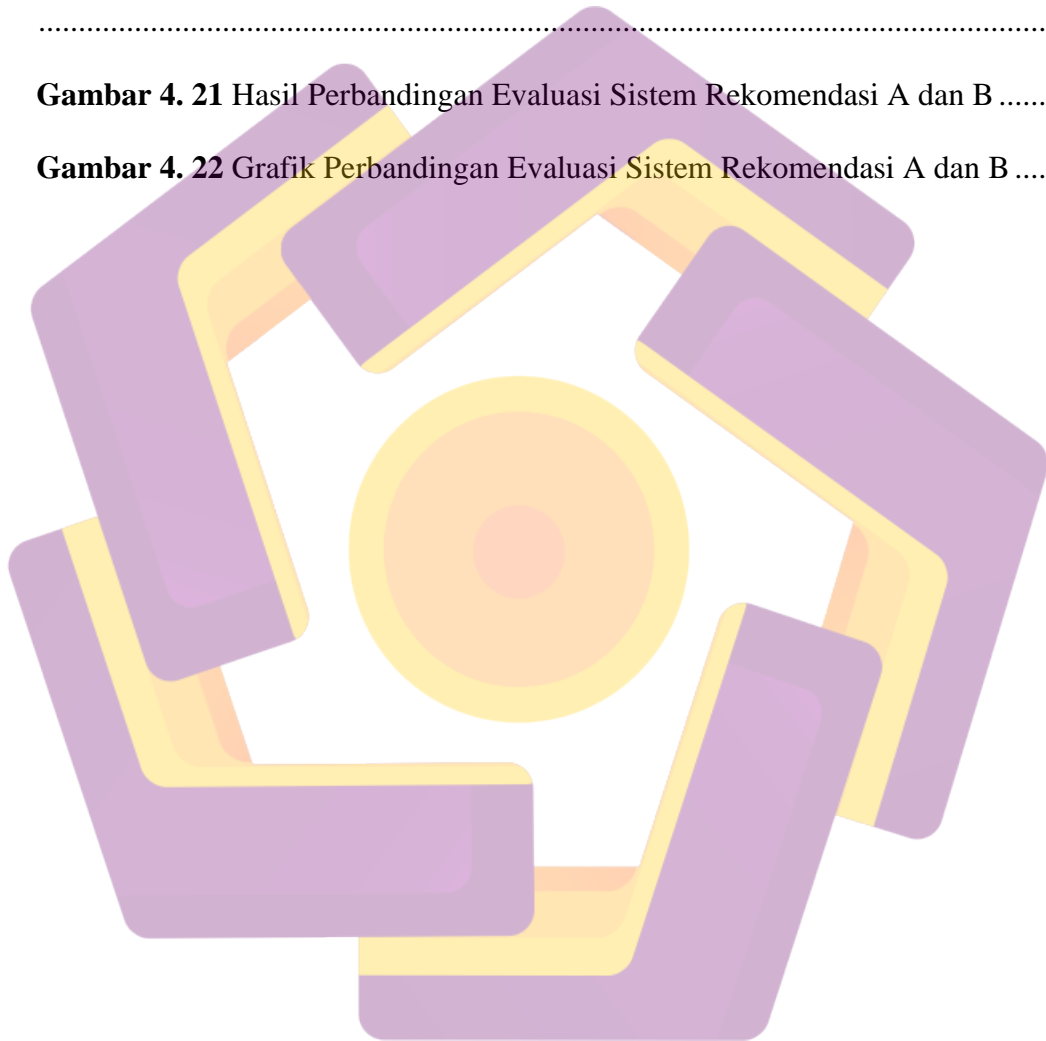
3.1.5.1.5	Menentukan Nilai N dan Mengambil Data Buku Berdasarkan Nilai N	49
3.1.6	Evaluasi.....	51
3.1.6.1	MAE.....	51
3.1.6.2	MSE.....	52
3.1.6.3	RMSE.....	53
3.2	Alat dan Bahan Penelitian.....	54
3.2.1	Kebutuhan Perangkat keras.....	54
3.2.2	Kebutuhan Perangkat lunak	54
BAB IV	HASIL DAN PEMBAHASAN	56
4.1	Pengambilan Data	56
4.2	Pemilihan Data.....	58
4.3	Pembersihan Data.....	58
4.4	Analisis Data	59
4.4.1	Analisis Data Sistem Rekomendasi A	59
4.4.2	Analisis Data Sistem Rekomendasi B.....	60
4.5	Pengolahan Data.....	61
4.5.1	Menentukan User X dan User Y	62
4.5.2	Proses K-Nearest Neighbor.....	63
4.5.2.1	Perhitungan Euclidean Distance	63

4.5.2.2	Menghitung Bobot User Y dan Menentukan Nilai K.....	65
4.5.3	Menghitung Nilai Rekomendasi Buku.....	66
4.5.4	Menghitung Nilai Predeksi Rating Buku.....	68
4.5.5	Pemilihan Rekomendasi Buku User X.....	68
4.6	Evaluasi.....	70
4.6.1	Hasil Evaluasi Sistem Rekomendasi A.....	72
4.6.1.1	MAE.....	72
4.6.1.2	MSE.....	73
4.6.1.3	RMSE.....	73
4.6.2	Hasil Evaluasi Sistem Rekomendasi B.....	75
4.6.2.1	MAE.....	75
4.6.2.2	MSE.....	75
4.6.2.3	RMSE.....	76
4.7	Hasil Perbandingan Evaluasi Sistem Rekomendasi A dan B.....	77
BAB V KESIMPULAN DAN SARAN.....		79
5.1	Kesimpulan.....	79
5.2	Saran.....	80
DAFTAR PUSTAKA.....		lxxxii

DAFTAR GAMBAR

Gambar 2. 1 Konsep User-Based Collaborative Filtering.....	21
Gambar 2. 2 Konsep Item-Based Collaborative Filtering	22
Gambar 3. 1 Diagram Alir Penelitian A	30
Gambar 3. 2 Diagram Alir Penelitian B	31
Gambar 3. 3 Diagram Alir Pengolahan Data.....	34
Gambar 4. 1 Sampel Data BX-Users.csv	56
Gambar 4. 2 Sampel Data BX-Books.csv	57
Gambar 4. 3 Sampel Data BX-Book-Ratings.csv	58
Gambar 4. 4 Total Data Rating Sistem Rekomendasi A	59
Gambar 4. 5 Total Data User yang Merating Sistem Rekomendasi A.....	59
Gambar 4. 6 Total Data Buku yang Dirating Sistem Rekomendasi A.....	59
Gambar 4. 7 Distribusi Rating.....	60
Gambar 4. 8 Total Data Rating Sistem Rekomendasi B	60
Gambar 4. 9 Total Data User yang Merating Sistem Rekomendasi B	60
Gambar 4. 10 Total Data Buku yang Dirating Sistem Rekomendasi B	61
Gambar 4. 11 Distribusi Rating Sistem Rekomendasi B	61
Gambar 4. 12 Hasil Rekomendasi Buku User X.....	70
Gambar 4. 13 Hasil Evaluasi MAE Sistem Rekomendasi A.....	72
Gambar 4. 14 Hasil Evaluasi MSE Sistem Rekomendasi A	73
Gambar 4. 15 Hasil Evaluasi RMSE Sistem Rekomendasi A.....	74
Gambar 4. 16 Grafik Perbandingan Peningkatan Evaluasi Sistem Rekomendasi A	74

Gambar 4. 17 Hasil Evaluasi MAE Sistem Rekomendasi B	75
Gambar 4. 18 Hasil Evaluasi MSE Sistem Rekomendasi B	76
Gambar 4. 19 Hasil Evaluasi RMSE Sistem Rekomendasi B	76
Gambar 4. 20 Grafik Perbandingan Peningkatan Evaluasi Sistem Rekomendasi B	77
Gambar 4. 21 Hasil Perbandingan Evaluasi Sistem Rekomendasi A dan B	77
Gambar 4. 22 Grafik Perbandingan Evaluasi Sistem Rekomendasi A dan B	78



DAFTAR TABEL

Tabel 3. 1 Sampel Data Rating.....	36
Tabel 3. 2 Rating Buku User X	37
Tabel 3. 3 Rating Buku User Y	37
Tabel 3. 4 Pengelompokan Rating User X	38
Tabel 3. 5 Pengelompokan Rating User Y	39
Tabel 3. 6 Nilai Similarity User Y dengan User X	40
Tabel 3. 7 Hasil Nilai Bobot User Y	42
Tabel 3. 8 Hasil K-NN dengan K=3.....	43
Tabel 3. 9 Nilai Rekomendasi Buku	45
Tabel 3. 10 Nilai Prediksi Rating Buku	47
Tabel 3. 11 Data Perangkingan Prediksi Rating Buku.....	49
Tabel 3. 12 Data Rekomendasi Buku User X.....	50
Tabel 3. 13 Rekomendasi Buku User X	50
Tabel 3. 14 Hasil Evaluasi MAE.....	51
Tabel 3. 15 Hasil Evaluasi MSE.....	52
Tabel 3. 16 Hasil Evaluasi RMSE.....	53
Tabel 4. 1 Perbandingan Hasil Analisa Sistem Rekomendasi A dan B	61

INTISARI

United Nations Educational Scientific, and Cultural Organization (UNESCO) mencatat minat baca masyarakat di Indonesia hanya mencapai 0,001% artinya dalam 1000 orang hanya satu orang yang mempunyai minat baca yang tinggi. Rendahnya minat baca di Indonesia tidak hanya dipengaruhi oleh ketidaktertarikan dalam kegiatan membaca tetapi juga dipengaruhi karena tidak mendapatkan buku yang tepat sesuai keinginan pembaca. Salah satu upaya yang dilakukan untuk meningkatkan minat baca adalah membangun sistem rekomendasi yang bertujuan untuk mendapatkan buku yang sesuai dengan keinginan pembaca.

Penelitian ini membandingkan dua sistem rekomendasi, yaitu sistem rekomendasi yang tidak memasuki tahap pembersihan data dan sistem rekomendasi yang memasuki tahap pembersihan data. Kedua sistem rekomendasi tersebut dibangun menggunakan teknik *user-based collaborative filtering* dengan algoritma *k-nearest neighbor*. *User-based collaborative filtering* merupakan teknik rekomendasi buku yang diberikan berdasarkan pertimbangan penilaian user sebelumnya. Nilai k pada algoritma *k-nearest neighbor* merupakan banyaknya *neighbor* yang digunakan untuk pertimbangan rekomendasi. Penelitian ini menggunakan beberapa nilai k pada kedua sistem rekomendasi, kemudian dilakukan evaluasi pada nilai k untuk menemukan nilai k terbaik yang akan digunakan pada sistem rekomendasi. Evaluasi menggunakan metode *mean absolute error* (MAE), *mean squared error* (MSE) dan *root mean squared error* (RMSE).

Hasil evaluasi menunjukkan bahwa nilai $k=1$ merupakan nilai k terbaik karena menghasilkan nilai error paling kecil pada kedua sistem rekomendasi. Sistem rekomendasi yang tidak memasuki tahap pembersihan data menghasilkan nilai error 0.011742 dengan metode MAE, 0.018774 dengan metode MSE dan 0.011742 dengan metode RMSE. Sistem rekomendasi yang memasuki tahap pembersihan data menghasilkan nilai error 0.011727 dengan metode MAE, 0.018117 dengan metode MSE dan 0.011727 dengan metode RMSE. Oleh karena itu, sistem rekomendasi yang memasuki tahap pembersihan data dengan nilai $k=1$ adalah sistem rekomendasi yang baik digunakan untuk merekomendasikan buku.

Kata Kunci: *Sistem Rekomendasi, User Based Collaborative Filtering, K-Nearest Neighbor, Euclidean Distance, MAE, MSE, RMSE.*

ABSTRACT

United Nations Educational Scientific, and Cultural Organization (UNESCO) realized that the reading awareness of the Indonesian people just only 0.001% it's means only one out of the 1000 people was concerned with reading awareness. The low literacy rate in Indonesia is not only caused by the people reading indolence but also difficult to find the book based on preference. The one thing to solving this issue is building a system that can recommend the book to the reader based on reader preference.

This research is comparing two methods of recommender system, those are recommender system without data cleaning process, and another one is using data cleaning process. Those methods were built using the same technique called user-based collaborative filtering with a k-nearest neighbor. User-based collaborative filtering is a technique to make the book recommendation based on previous user rating data. K value in K-nearest neighbor reflects the amount of the neighbor which is used to consider the recommendation result. This research uses several k values on both methods, then we are evaluating to find which best k value is used on our recommender system. The evaluations are using mean absolute error (MAE), mean squared error (MSE) and root mean squared error (RMSE).

The evaluation results showed that $k=1$ is the best k value because producing the smallest error value on both recommender system methods. The recommender system without the data cleaning process got an error value 0.011742 on MAE, 0.018774 on MSE, and 0.011742 on RMSE. On another side, the recommender system using the data cleaning process got an error value 0.011727 on MAE, 0.018117 on MSE, and 0.011727 on RMSE. Based on those evaluation results, the Recommendation System with data cleaning and $k=1$ is the best method to make a book recommendation system.

Keywords : Recommender System, User Based Collaborative Filtering, K-Nearest Neighbor, Euclidean Distance, MAE, MSE, RMSE.