

**PENGARUH USER WEIGHTING PADA NEIGHBOURHOOD BASED  
COLLABOARTIVE FILTERING MENGGUNAKAN FIREFLY  
ALGORITHM**

**SKRIPSI**



disusun oleh

**Yahya Sugandl.Y**

**17.11.1604**

**PROGRAM SARJANA  
PROGRAM STUDI INFORMATIKA  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2020**

**PENGARUH USER WEIGHTING PADA NEIGHBOURHOOD BASED  
COLLABOARTIVE FILTERING MENGGUNAKAN FIREFLY  
ALGORITHM**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Informatika



disusun oleh

**Yahya Sugandl.Y**

**17.11.1664**

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

**2020**

## PERSETUJUAN

### SKRIPSI

#### PENGARUH USER WEIGHTING PADA NEIGHBOURHOOD BASED COLLABOARTIVE FILTERING MENGGUNAKAN FIREFLY ALGORITHM

yang dipersiapkan dan disusun oleh

**Yahya SugandI.Y**

17.11.1664

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 10 April 2020

**Dosen Pembimbing,**

Hartatik, ST., M.Cs.  
NIK. 190302232

## PENGESAHAN

### SKRIPSI

#### PENGARUH USER WEIGHTING PADA NEIGHBOURHOOD BASED COLLABOARTIVE FILTERING MENGGUNAKAN FIREFLY ALGORITHM

yang dipersiapkan dan disusun oleh

**Yahya Sugandi.Y**

**17.11.1664**

telah dipertahankan di depan Dewan Penguji  
pada tanggal 18 November 2020

#### Susunan Dewan Penguji

**Nama Penguji**

**Tanda Tangan**

**Ninik Tri Hartanti, M.Kom**

**NIK. 190302330**

**Alfie Nur Rahmi, M.Kom**

**NIK. 190302240**

**Hartatik, ST., M.Cs.**

**NIK. 190302232**

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 3 Desember 2020

**DEKAN FAKULTAS ILMU KOMPUTER**

**Krisnawati, S.Si, M.T.**

**NIK. 190302038**

## 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, 2 Desember 2020



Yahya Sugandi.Y

NIM. 17.11.1664

## MOTTO

“Hiduplah seolah kamu akan mati besok.  
Belajarlah seolah kamu akan hidup selamanya.”  
-Mahatma Gandhi

“Bermimpilah dalam hidup, jangan hidup dalam mimpi”  
-Andrea Hirata

“Tidak masalah seberapa lambat kau berjalan asalkan kau tidak berhenti”  
-Confucius

“Perbuatlah segala sesuatu atas kehendak mu sendiri. Agar jika kau menyesalinya,  
kau tidak akan menyalahkan orang lain”  
- Yahya Sugandi.Y

“Kuatkan dan teguhkanlah hatimu, janganlah takut dan jangan gemetar karena mereka, sebab TUHAN, Allahmu, Dialah yang berjalan menyertai engkau; Ia tidak akan membiarkan engkau dan tidak akan meninggalkan engkau.”  
-Ulangan 31:16

## PERSEMBAHAN

Puji syukur saya panjatkan kehadiran Tuhan Yesus Kristus yang telah memberikan berkat yang luar biasa kepada saya, sehingga saya bisa menyelesaikan skripsi ini dengan baik. Saya juga sangat berterima kasih kepada orang-orang yang telah secara langsung maupun tidak langsung yang telah membantu saya dalam menyelesaikan skripsi ini. Skripsi ini saya persembahkan kepada:

1. Bapak dan Ibu yang tidak pernah lelah untuk mendoakan, mendukung dan memberi nasihat untuk menjadi pribadi yang lebih baik.
2. Keluarga besar yang selalu mendoakan, memberi semangat dan nasehat kepada saya.
3. Kakak dan Adik yang selalu men *support* dan memberikan motivasi.
4. Bayu Permana Sejati, sebagai rekan seperjuangan dalam penelitian *recommender system* yang selalu memberikan dorongan dan masukan untuk terus berjuang sampai akhir.
5. Team Bunglon, Rama dan Rengga yang telah banyak membantu, membagi ilmu dan masukan kepada saya.
6. Teman – teman seperjuangan yang melakukan penelitian, Adit, Aan, Nisah, Nisah, Visq, Ibam, Nurma, and Christy, Arya, Erick, Anang, Fayami, Nur Hidayah, Wahidin, and Dimas.
7. Teman – teman kontrakan, Agung, Hamdi, Dio, Jo, Agil, Lucas, Dok, Frans, Adi Sanjaya, Adi Tugel.
8. Seluruh pihak yang tidak saya sebutkan satu per satu, terimakasih atas segala bantuan dan doanya sehingga terselesaikan skripsi ini.

## KATA PENGANTAR

Puji dan syukur saya panjatkan Tuhan Yesus yang telah memberikan kasih, kekuatan, dan sukacita sehingga saya dapat menyelesaikan skripsi yang berjudul Pengaruh *User Weighting* pada *Neighbourhood Based Collaborative Filtering* menggunakan *Firefly Algorithm*.

Skripsi ini saya buat guna menyelesaikan studi jenjang Strata Satu (S1) pada program studi Informatika fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta. Selain itu juga merupakan suatu bukti bahwa mahasiswa telah menyelesaikan kuliah jenjang program strata satu dan untuk memperoleh gelar Sarjana Komputer. Dengan selesainya skripsi ini, maka pada kesempatan ini saya mengucapkan terima kasih kepada :

1. Bapak Prof. Dr. M. Suyanto, MM. selaku Rektor Universitas AMIKOM Yogyakarta.
2. Ibu Krisnawati, S.Si, M.T. selaku Dekan Fakultas Ilmu Komputer Universitas AMIKOM Yogyakarta.
3. Bapak Sudarmawan, S.T., M.T. selaku Ketua Program Studi SI Informatika Universitas AMIKOM Yogyakarta.
4. Ibu Hartatik, ST., M.Cs. selaku dosen pembimbing yang selalu bijaksana memberikan bimbingan nasehat serta waktu nya selama penulisan skripsi ini.
5. Dosen Penguji, Ibu Ninik Tri Hartanti, M.Kom, Ibu Alfie Nur Rahmi, M.Kom, dan segenap Dosen serta Karyawan Universitas AMIKOM Yogyakarta yang telah berbagi ilmu dan pengalamannya.



6. Bapak Joko Dwi Santoso, M.Kom yang sudah berbagi ilmu dan pengalamannya, serta memberikan nasihat untuk menjadi lebih baik.
7. Teman – teman pengurus Forum Asisten Universitas AMIKOM Yogyakarta periode 2019/2020 yang telah memberikan motivasi dan dorongan kepada saya.
8. Pengurus Forum Asisten Universitas AMIKOM Yogyakarta periode 2018/2019 dan 2020/2021 yang telah berbagi ilmu dan pengalamannya.
9. Teman - teman 17 Informatika 11 yang sudah berjuang bersama – sama, membagi ilmu dan kebersamaannya selama perkuliahan.
10. Seluruh warga Asrama TPN Yogyakarta, yang bersama – sama berjuang menempuh pendidikan.
11. Teman – teman Amikom Overclocking, yang sudah memberikan masukan dan wawasan kepada saya.
12. Semua pihak yang telah membantu baik secara langsung maupun tidak langsung.

Semoga Tuhan memberikan balasan yang lebih kepada semua yang telah ikut membantu saya hingga menyelesaikan skripsi ini. Demi perbaikan selanjutnya, saran dan kritik yang membangun diterima dengan senang hati dan rasa terima kasih. Semoga skripsi ini dapat bermanfaat bagi sayadan kita semua.

Yogyakarta, 2 Desember 2020



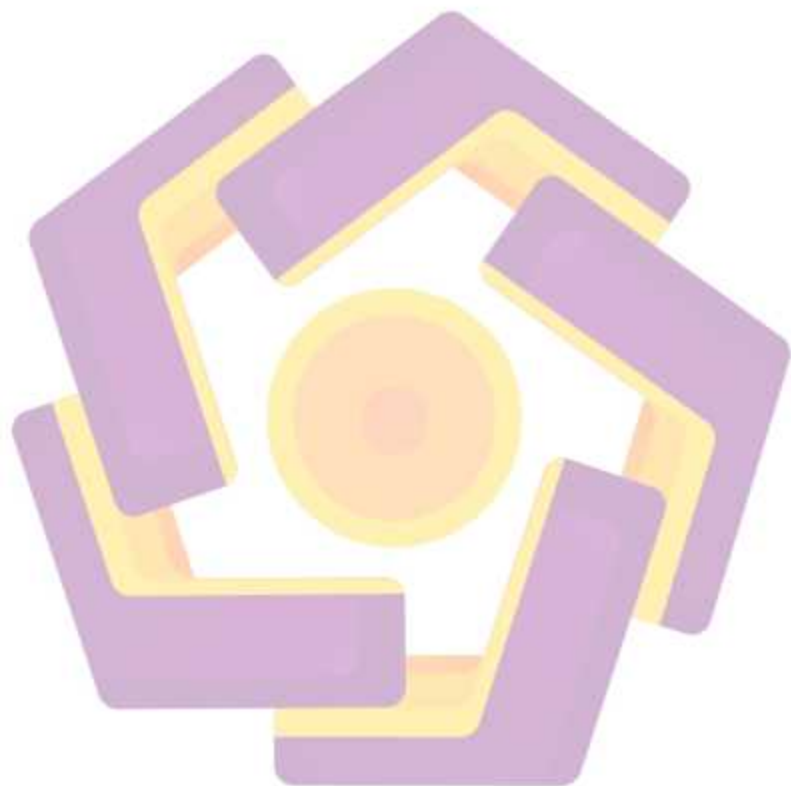
Yahya Sugandi.Y

## DAFTAR ISI

JUDUL .....	I
PERSETUJUAN.....	II
PENGESAHAN.....	III
PERNYATAAN.....	IV
MOTTO .....	V
PERSEMBAHAN.....	VI
KATA PENGANTAR.....	VII
DAFTAR ISI.....	IX
DAFTAR TABEL .....	XIII
DAFTAR GAMBAR.....	XIV
DAFTAR ISTILAH .....	XVI
INTISARI .....	XVII
ABSTRACT.....	XVIII
<b>BAB I PENDAHULUAN.....</b>	<b>1</b>
1.1 LATAR BELAKANG.....	1
1.2 RUMUSAN MASALAH.....	5
1.3 BATASAN MASALAH.....	5
1.4 MAKSUD DAN TUJUAN PENELITIAN.....	6
1.5 MANFAAT PENELITIAN.....	6
1.6 METODOLOGI PENELITIAN .....	6
1.6.1 Studi Literatur.....	6
1.6.2 Analisis Permasalahan .....	6
1.6.3 Perancangan Sistem .....	7
1.6.4 Implementasi.....	7
1.6.5 Pengujian.....	7

1.6.6	Penyusunan Laporan .....	7
1.7	SISTEMATIKA PENULISAN .....	7
<b>BAB II LANDASAN TEORI .....</b>		<b>9</b>
2.1	TINJAUAN PUSTAKA.....	9
2.2	SISTEM REKOMENDASI.....	14
2.2.1	Content-Based Filtering .....	14
2.2.2	Collaborative Filtering .....	15
2.2.3	Alur Penerapan Metode Collaborative Filtering .....	20
2.3	UTILITY MATRIX.....	20
2.4	PEARSON CORRELATION COEFFICIENT.....	21
2.5	COSINE SIMILARITY .....	22
2.6	SWARM INTELLIGENCE.....	23
2.6.1	Firefly Algorithm .....	25
2.6.2	Alur Firefly Algorithm.....	27
2.7	K-NEAREST NEIGHBOR.....	27
2.8	EXPLORATORY DATA ANALYSIS.....	28
2.9	CROSS VALIDATION.....	28
2.10	PERHITUNGAN PREDIKSI.....	29
2.11	PENGUKURAN EVALUASI.....	30
<b>BAB III METODE PENELITIAN .....</b>		<b>32</b>
3.1	ALAT DAN BAHAN PENELITIAN.....	32
3.1.1	Kebutuhan Perangkat Keras.....	32
3.1.2	Kebutuhan Perangkat Lunak.....	32
3.2	SAMPel DATA.....	32
3.3	EXPLORATORY DATA ANALYSIS.....	33
3.3.1	Jumlah Film Yang Dirilis Setiap Tahun.....	34
3.3.2	Pengguna Yang Banyak Merating Film.....	35
3.3.3	Jumlah Pengguna Berdasarkan Pekerjaan.....	36
3.3.4	Persebaran Jenis Kelamin Pengguna.....	36
3.3.5	Persebaran Rating Pada Dataset.....	37

3.4	PIPELINE PENELITIAN .....	38
3.5	PREPROCESSING .....	40
3.6	TRANSFORMASI DATA.....	40
3.6.1	Utility Matrix .....	40
3.6.2	Similarity Matrix.....	42
3.7	PEMBOBOTAN .....	43
3.8	TRAINING.....	44
3.9	TESTING .....	45
3.10	ANALISIS HASIL.....	46
<b>BAB IV IMPLEMENTASI DAN PENGUJIAN .....</b>		<b>47</b>
4.1	IMPLEMENTASI DATA.....	47
4.2	PRE-PROCESSING DATA.....	50
4.3	TRANSFORMASI DATA.....	51
4.3.1	Transformasi Utility Matrix .....	51
4.3.2	Transformasi Similarity Matrix.....	52
4.4	PEMBOBOTAN .....	53
4.5	TRAINING.....	54
4.6	TESTING.....	55
4.7	EKSPERIMEN DAN HASIL.....	55
4.7.1	Pengaruh Jumlah Pengguna .....	55
4.7.2	Pengaruh Jumlah K.....	60
4.7.3	Pengaruh Jumlah Data Training.....	64
4.7.4	Pengaruh User - based dan Item - based .....	68
4.8	HASIL REKOMENDASI.....	71
4.8.1	Rekomendasi Metode Tradisional.....	71
4.8.2	Rekomendasi Metode Tradisional Dan FA.....	73
<b>BAB V PENUTUP.....</b>		<b>76</b>
5.1	KESIMPULAN.....	76
5.2	SARAN.....	77



## DAFTAR TABEL

Tabel 2. 1 Tabel Penelitian Terdahulu .....	12
Tabel 3. 1 Distribusi dataset MovieLens 1M .....	33
Tabel 3. 2 Bentuk data awal .....	41
Tabel 3. 3 Bentuk data setelah ditransformasi menjadi utility matrix .....	41
Tabel 3. 4 Bentuk transformasi data menjadi similarity matrix .....	42
Tabel 4. 1 Hasil eksperimen jumlah 100 pengguna .....	56
Tabel 4. 2 Hasil eksperimen jumlah 200 pengguna .....	56
Tabel 4. 3 Hasil eksperimen jumlah 300 pengguna .....	57
Tabel 4. 4 Hasil eksperimen jumlah 6040 pengguna .....	57
Tabel 4. 5 Hasil eksperimen jumlah k pada 100 pengguna .....	61
Tabel 4. 6 Hasil eksperimen jumlah pada 200 pengguna .....	61
Tabel 4. 7 Hasil eksperimen jumlah k pada 300 pengguna .....	61
Tabel 4. 8 Hasil eksperimen jumlah k pada 6040 pengguna .....	62
Tabel 4. 9 Hasil MAE dari jumlah data training .....	64
Tabel 4. 10 Hasil RMSE dari jumlah data train .....	65
Tabel 4. 11 Hasil waktu eksekusi dari jumlah data train .....	65
Tabel 4. 12 Hasil MAE user-based dan item-based .....	68
Tabel 4. 13 Hasil RMSE user-based dan item-based .....	68
Tabel 4. 14 Hasil Waktu eksekusi user-based dan item-based .....	69
Tabel 4. 15 Hasil rekomendasi metode tradisional dengan 100 pengguna .....	71
Tabel 4. 16 Hasil rekomendasi metode tradisional dengan 200 pengguna .....	72
Tabel 4. 17 Hasil rekomendasi metode tradisional dengan 300 pengguna .....	72
Tabel 4. 18 Hasil rekomendasi metode tradisional dengan 6040 pengguna .....	73
Tabel 4. 19 Hasil rekomendasi metode tradisional dan fa dengan 100 pengguna	73
Tabel 4. 20 Hasil rekomendasi metode tradisional dan fa dengan 200 pengguna	74
Tabel 4. 21 Hasil rekomendasi metode tradisional dan fa dengan 300 pengguna	75
Tabel 4. 22 Hasil rekomendasi metode tradisional dan fa dengan 6040 pengguna	75
.....	75

## DAFTAR GAMBAR

Gambar 1. 1 Perbandingan pendapatan retail pada tahun 1990 dan 2021 .....	3
Gambar 2. 1 Skema sistem rekomendasi dengan metode content based .....	15
Gambar 2. 2 Skema sistem rekomendasi dengan metode collaborative filtering .	16
Gambar 2. 3 Utility matix yang merepresentasikan ratings dari film dengan skala 1-5 .....	21
Gambar 2. 4 Konsep Swarm Intelligence .....	24
Gambar 2. 5 Ilustrasi Cross Validation .....	29
Gambar 3. 1 Jumlah film yang dirilis setiap tahun .....	34
Gambar 3. 2 Pengguna yang banyak merating film .....	35
Gambar 3. 3 Jumlah pengguna berdasarkan pekerjaan .....	36
Gambar 3. 4 Persebaran jenis kelamin pengguna .....	37
Gambar 3. 5 Persebaran rating pada dataset .....	37
Gambar 3. 6 Pipeline Penelitian .....	39
Gambar 3. 7 Proses Training .....	44
Gambar 3. 8 Proses Testing .....	45
Gambar 4. 1 Search dataset melalui browser .....	47
Gambar 4. 2 Tampilan situs grouplens.org untuk mengambil data .....	47
Gambar 4. 3 Direktori penyimpanan data .....	48
Gambar 4. 4 Tampilan direktori pada jupyter notebook .....	48
Gambar 4. 5 Sampel data .....	49
Gambar 4. 6 Proses ekstraksi fitur .....	50
Gambar 4. 7 Data hasil ekstraksi fitur .....	50
Gambar 4. 8 Transformasi utility matrix .....	51
Gambar 4. 9 Hasil utility matrix .....	51
Gambar 4. 10 Transformasi similarity matrix .....	52
Gambar 4. 11 Hasil similarity matrix .....	52
Gambar 4. 12 Inisiasi parameter firefly algorithm .....	53
Gambar 4. 13 Hasil pembobotan atau ranking pengguna .....	53
Gambar 4. 14 Proses pengambilan pengguna aktif .....	54

Gambar 4. 15 Proses training model KNN .....	54
Gambar 4. 16 Proses testing model KNN .....	55
Gambar 4. 17 Perbandingan nilai MAE 100, 200, 300,6040 pengguna .....	58
Gambar 4. 18 Perbandingan nilai RMSE 100, 200, 300, 6040 pengguna .....	58
Gambar 4. 19 Perbandingan waktu eksekusi 100, 200, 300, 6040 pengguna.....	59
Gambar 4. 20 Perbandingan nilai MAE k = 10, 20, 30, 40, 50.....	62
Gambar 4. 21 Perbandingan nilai RMSE k = 10, 20, 30, 40, 50.....	63
Gambar 4. 22 Perbandingan waktu eksekusi k = 10, 20, 30, 40, 50.....	63
Gambar 4. 23 Perbandingan nilai RMSE dari jumlah data training .....	67
Gambar 4. 24 Perbandingan waktu eksekusi dari jumlah data training.....	67
Gambar 4. 25 Perbandingan nilai MAE dari metode CF.....	69
Gambar 4. 26 Perbandingan nilai RMSE dari metode CF.....	70
Gambar 4. 27 Perbandingan waktu eksekusi dari metode CF .....	70





## DAFTAR ISTILAH

- FA Pearson : Sejumlah pengguna yang telah dibobotkan menggunakan *firefly algorithm* dengan pencarian kesamaan *pearson correlation* yang digunakan untuk data *training* dan *testing*.
- FA Cosine : Sejumlah pengguna yang telah dibobotkan menggunakan *firefly algorithm* dengan pencarian kesamaan *cosine similarity* yang digunakan untuk data *training* dan *testing*.
- Similarity : Nilai kesamaan antar pengguna yang diukur dengan metode pencarian kesamaan.
- Metode Tradisional : Pada sistem rekomendasi, metode tradisional adalah metode untuk mencari kesamaan antar pengguna atau item.
- Pengguna Aktif : Adalah pengguna yang memiliki *similarity* tertinggi pada dataset dan didapatkan dari proses optimasi menggunakan *firefly algorithm*.
- Sparsity : Kondisi tidak memiliki cukup sesuatu atau jarang data dalam sebuah model data.
- Feature : didalam dataset terdapat kolom dan row, *feature* adalah istilah untuk sebuah kolom.

## INTISARI

Sistem Rekomendasi merupakan sebuah sistem yang mampu memberikan konten yang berkemungkinan besar akan dipilih oleh pengguna karena konten yang diberikan akan berdasarkan penyaringan informasi yang mengambil preferensi dari perilaku maupun riwayat pengguna. Baru – baru ini, para peneliti melakukan penelitian untuk meningkatkan kualitas rekomendasi salah satunya dengan menggunakan *swarm intelligence* pada sistem rekomendasi tradisional *collaborative filtering*.

Penelitian ini bertujuan untuk mengetahui pengaruh *user weighting* pada sistem rekomendasi tradisional salah satu *swarm intelligence* yaitu *firefly algorithm* untuk memberikan bobot pada pengguna dan mendapatkan pengguna aktif pada dataset.

Pada penelitian ini, *firefly algorithm* bisa mendapatkan 100 pengguna yang memiliki pengaruh besar pada hasil prediksi dengan nilai error MAE sebesar 0.8101. Namun menggunakan semua data bisa memberikan nilai error MAE yang lebih rendah yaitu sebesar 0.8007.

**Kata Kunci** : Sistem Rekomendasi, *Firefly Algorithm*, *Pearson Correlation Coefficient*, *Cosine Similarity*, Pembobotan Pengguna



## ABSTRACT

*The Recommendation System is a system that is able to provide content that is likely to be chosen by users because the content provided will be based on filtering information that takes preferences from user behavior and history. Recently, the researchers conducted research to improve the quality of one of the recommendations by using swarm intelligence in the traditional recommendation system collaborative filtering.*

*This study aims to determine the effect of user weighting on a traditional recommendation system, one of the swarm intelligence, namely firefly algorithm to give weight to users and get active users on the dataset.*

*In this study, the firefly algorithm was able to find 100 users who had a big influence on the prediction results with an MAE error value of 0.8101. However, using all the data can give a lower MAE error value of 0.8007.*

**Keywords** : Recommendation System, Firefly Algorithm, Pearson Correlation Coefficient, Cosine Similarity, User Weighting

