

**PEMODELAN COLLABORATIVE FILTERING MENGGUNAKAN  
NAIVE BAYES PADA DATASET MOVIELENS**

**SKRIPSI**



disusun oleh

**Rahmat Kurniawan**

**19.22.2232**

**PROGRAM SARJANA  
PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2021**

**PEMODELAN COLLABORATIVE FILTERING MENGGUNAKAN  
NAIVE BAYES PADA DATASET MOVIELENS**

**SKRIPSI**

untuk memenuhi sebagian persyaratan  
mencapai gelar Sarjana  
pada Program Studi Sistem Informasi



disusun oleh

**Rahmat Kurniawan**

**19.22.2232**

**PROGRAM SARJANA  
PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS AMIKOM YOGYAKARTA  
YOGYAKARTA  
2021**

# PERSETUJUAN

## SKRIPSI

### PEMODELAN COLLABORATIVE FILTERING MENGGUNAKAN NAIVE BAYES PADA DATASET MOVIELENS

yang dipersiapkan dan disusun oleh

**Rahmat Kurniawan**

**19.22.2232**

telah disetujui oleh Dosen Pembimbing Skripsi  
pada tanggal 1 September 2020

**Dosen Pembimbing,**

**Hartatik, S.T., M.Cs.**

**NIK. 190302232**

**PENGESAHAN**

**SKRIPSI**

**PEMODELAN COLLABORATIVE FILTERING MENGGUNAKAN  
NAIVE BAYES PADA DATASET MOVIELENS**

yang dipersiapkan dan disusun oleh

**Rahmat Kurniawan**

**19.22.2232**

telah dipertahankan di depan Dewan Penguji  
pada tanggal 19 Februari 2021

**Susunan Dewan Penguji**

**Nama Penguji**

**Tanda Tangan**

**Rizqi Sukma Kharisma, M.Kom**  
**NIK. 190302215**

**Acihmah Sidauruk, M.Kom**  
**NIK. 190302238**

**Hartatik, S.T., M.Cs.**  
**NIK. 190302232**

Skripsi ini telah diterima sebagai salah satu persyaratan  
untuk memperoleh gelar Sarjana Komputer  
Tanggal 19 Maret 2021

**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, 18 Maret 2021



Rahmat Kurniawan

NIM. 19.22.2232

## MOTTO

Setiap orang adalah guru, setiap rumah adalah sekolah

- Ki Hajar Dewantara -

*Fa inna ma 'al- 'usri yusrā, inna ma 'al- 'usri yusrā*

QS: Al-Insyirah, 5-6

*Work Hard, Play Hard*

- Wiz Khalifa -

*Yen kuat dilakoni, yen ra kuat ditinggal ngopi*

- Penulis -

## PERSEMBAHAN

Dengan segenap usaha dan ucapan syukur Penulis panjatkan kepada Allah SWT, karena-Nya lah Penulis dapat menyelesaikan Skripsi ini. Akhirnya skripsi ini penulis persembahkan untuk :

1. Kedua orang tua dan keluarga yang selalu mendukung dan memberi dukungan hingga akhirnya dapat menyelesaikan kuliah dengan baik.
2. Ibu Hartatik, S.T., M.Cs, selaku dosen pembimbing saya terimakasih telah memberikan arahan serta bimbingannya dari awal hingga akhir Skripsi ini.
3. Rekan kerja saya sekaligus kolega kampus dari Symbolic Bayuarga Damar Sungkowo, S.Kom, M Yusuf, S.Kom beserta rekan-rekan kerja lainnya terimakasih telah kebersamai saya dalam berproses selama ini.
4. Teman-teman 19-SITtransfer, terimakasih atas kebersamaannya selama ini.

## KATA PENGANTAR

**Assalamu'alaikum Wr. Wb.**

Puji syukur alhamdulillah saya panjatkan kehadiran Allah SWT atas limpahan rahmat dan inayah-Nya, sehingga penulis dapat menyelesaikan Skripsi ini dengan baik. Sholawat dan salam semoga terhaturkan kepada manusia pilihan dan suri tauladan terbaik, Nabi Muhammad SAW, yang telah menunjukkan jalan kebenaran kepada umat manusia.

Adapun Skripsi ini dibuat untuk memenuhi syarat guna memperoleh gelar Strata (S.Kom) Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Bagi penulis, proses penyusunan laporan Skripsi ini tidak mudah. Banyak kekurangan dan hambatan yang penulis alami dikarenakan keterbatasan kemampuan penulis sendiri. Penulis sadari ada banyak pihak yang ikut membantu dan memberi dukungan kepada penulis sehingga Skripsi ini dapat terselesaikan. Oleh karena itu, penulis ucapkan banyak terima kasih kepada semua pihak yang terlibat, terutama kepada :

1. Allah SWT atas limpahan rahmat, hidayah dan nikmat kehidupan.
2. Nabi Muhammad SAW sebagai Nabi dan suri tauladan bagi umat-Nya.
3. Orang tua dan seluruh keluarga tercinta yang selalu mendoakan dan memotivasi dengan tulus ikhlas sehingga menjadi pelecut bagi penulis untuk menyelesaikan Skripsi ini.
4. Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta.



5. Krisnawati, S.Si., M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta yang telah mendukung pembuatan Skripsi ini.
6. Hartatik, S.T., M.Cs, selaku pembimbing yang telah memberikan dukungan serta bimbingannya dalam menyusun Skripsi ini.
7. Bapak/Ibu dosen yang telah memberikan dukungannya dalam pembuatan laporan ini.
8. Serta teman-teman yang telah memberikan kritik dan saran yang sangat membantu dalam pembuatan Skripsi ini.

Akhir kata penyusun ucapkan terima kasih dan semoga laporan ini berguna bagi pembaca. Penyusun menyadari bahwa dalam Menyusun laporan Skripsi ini masih jauh dari sempurna. Untuk itu, penulis mohon maaf bila ada kesalahan dalam pembuatan laporan ini, sehingga penulis menerima kritik dan saran yang membangun dari pembaca.

Yogyakarta, 18 Maret 2021

Rahmat Kurniawan

## DAFTAR ISI

JUDUL .....	II
PERSETUJUAN .....	III
PENGESAHAN .....	IV
PERNYATAAN.....	V
MOTTO.....	VI
PERSEMBAHAN.....	VII
KATA PENGANTAR .....	VIII
DAFTAR ISI.....	X
DAFTAR TABEL.....	XIII
DAFTAR GAMBAR.....	XIV
DAFTAR PERSAMAAN MATEMATIKA.....	XVI
INTISARI.....	XVII
ABSTRACT.....	XVIII
BAB I PENDAHULUAN.....	1
1.1 LATAR BELAKANG.....	1
1.2 RUMUSAN MASALAH.....	2
1.3 BATASAN MASALAH.....	2
1.4 TUJUAN PENELITIAN.....	3
1.5 MANFAAT PENELITIAN.....	3
1.6 METODE PENELITIAN.....	3
1.6.1 Metode Pengumpulan Data.....	3
1.6.2 Metode Analisis.....	3
1.6.3 Metode Pengembangan Sistem.....	4
1.7 SISTEMATIKA PENULISAN.....	5
BAB II LANDASAN TEORI.....	6
2.1 KAJIAN PUSTAKA.....	6
2.2 TEORI DAN KONSEP SISTEM PAKAR.....	8
2.2.1 Sistem Pakar.....	8
2.2.2 Manfaat dan Kekurangan Sistem Pakar.....	9

2.2.3	Ciri-Ciri Sistem Pakar .....	9
2.2.4	Klasifikasi Sistem Pakar .....	10
2.2.5	Konsep Dasar Sistem Pakar .....	11
2.2.6	Komponen Sistem Pakar .....	12
2.3	SISTEM REKOMENDASI.....	12
2.4	ALGORITMA .....	12
3.4.1	Pengertian Algoritma .....	12
3.4.2	Struktur Dasar Algoritma .....	13
2.5	COLLABORATIVE FILTERING .....	13
2.6	NAIVE BAYES.....	14
2.7	FLOWCHART.....	15
2.8	BAHASA PEMROGRAMAN YANG DIGUNAKAN.....	16
2.8.1	HTML .....	16
2.8.2	PHP .....	16
2.9	CODEIGNITER FRAMEWORK .....	16
2.10	MEAN SQUARED ERROR.....	17
2.11	PERANGKAT LUNAK YANG DIGUNAKAN .....	18
2.11.1	Visual Studio Code .....	18
2.11.2	Web Browser.....	19
2.11.3	XAMPP .....	19
2.11.4	Microsoft Excel.....	21
2.11.5	Bash Shell.....	21
2.11.6	Draw.io.....	22
<b>BAB III ANALISIS DAN PERANCANGAN .....</b>		<b>23</b>
3.1	ANALISIS DATASET .....	23
3.2	ANALISIS KEBUTUHAN SISTEM .....	24
3.2.1	Analisis Kebutuhan Fungsional .....	24
3.2.2	Analisis Kebutuhan Non Fungsional .....	25
3.3	ANALISIS PEMODELAN .....	26
3.3.1	Desain Metode .....	26
3.3.2	Formulasi Metode .....	27

3.3.3	Penghitungan Prediksi.....	29
3.3.4	Penghitungan Reliabilitas.....	29
3.4	PERANCANGAN SISTEM.....	30
3.4.1	Perancangan Pipeline .....	30
3.4.2	Perancangan Flowchart .....	32
3.4.3	Perancangan Database.....	32
3.4.4	Perancangan Antar Muka.....	36
<b>BAB IV IMPLEMENTASI DAN PEMBAHASAN .....</b>		<b>50</b>
1.1	DATABASE DAN TABEL .....	50
1.2	IMPLEMENTASI ANTAR MUKA .....	53
1.2.1	Halaman Login.....	53
1.2.2	Halaman Dashboard.....	53
1.2.3	Halaman Admin .....	55
1.2.4	Halaman User.....	55
1.2.5	Halaman Occupation.....	56
1.2.6	Halaman Item (Movie).....	56
1.2.7	Halaman Data (Rating) .....	57
1.2.8	Halaman Rekomendasi (Running Example).....	57
1.3	ANALISIS DAN PENGUJIAN .....	59
1.3.1	Sampel Perhitungan Manual .....	59
1.3.2	Penjelasan Rekomendasi.....	64
1.3.3	Hasil Pengujian .....	64
1.4	PEMBAHASAN LISTING PROGRAM.....	66
<b>BAB V PENUTUP.....</b>		<b>69</b>
5.1	KESIMPULAN.....	69
5.2	SARAN.....	69
<b>DAFTAR PUSTAKA .....</b>		<b>70</b>

## DAFTAR TABEL

Tabel 2.1 Perbandingan dengan Penelitian Sebelumnya .....	7
Tabel 2.2 Daftar Simbol Document Flowchart .....	15
Tabel 3.1 Main Properties .....	24
Tabel 3.2 Tabel Admin .....	33
Tabel 3.3 Tabel User .....	33
Tabel 3.4 Tabel Occupation .....	33
Tabel 3.5 Tabel Item .....	34
Tabel 3.6 Tabel Data .....	35
Tabel 4.1 Rating Matrix .....	60
Tabel 4.2 Prior .....	61
Tabel 4.3 Likelihood .....	61
Tabel 4.4 Classification Score .....	62
Tabel 4.5 Hasil Rekomendasi Dengan 5 Neighbor .....	63
Tabel 4.6 Hasil Rekomendasi Dengan 10 Neighbor .....	64
Tabel 4.7 Jumlah Error Kuadrat Data Training .....	65
Tabel 4.8 Tabel Mse .....	65

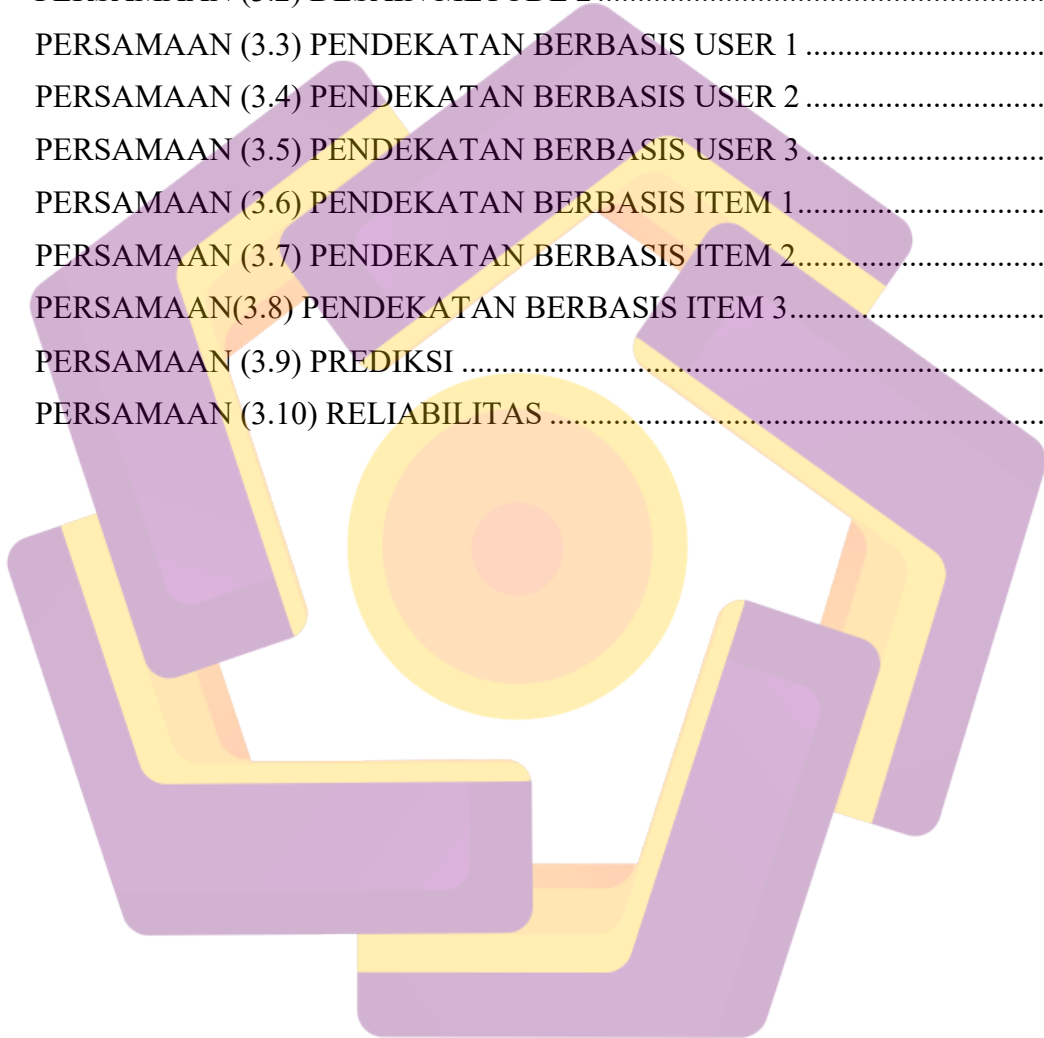
## DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Cara Kerja Mvc .....	17
Gambar 2.2 Tampilan Aplikasi Visual Studio Code.....	19
Gambar 2.3 Tampilan Aplikasi Xampp .....	20
Gambar 2.4 Tampilan Aplikasi Microsoft Excel .....	21
Gambar 2.5 Tampilan Aplikasi Bash Shell.....	22
Gambar 2.6 Tampilan Aplikasi Draw.io.....	22
Gambar 3.1 Melihat Dataset Dengan Terminal .....	24
Gambar 3.2 Pipeline.....	31
Gambar 3.3 Flowchart.....	32
Gambar 3.4 Relasi Tabel.....	35
Gambar 3.5 Rancangan Tampilan Login .....	36
Gambar 3.6 Rancangan Tampilan Dashboard .....	37
Gambar 3.7 Rancangan Tampilan Admin.....	38
Gambar 3.8 Rancangan Tampilan Tambah Admin.....	38
Gambar 3.9 Rancangan Tampilan Ubah Admin.....	39
Gambar 3.10 Rancangan Tampilan User .....	40
Gambar 3.11 Rancangan Tampilan Tambah User .....	40
Gambar 3.12 Rancangan Tampilan Ubah User.....	41
Gambar 3.13 Rancangan Tampilan Occupation .....	42
Gambar 3.14 Rancangan Tampilan Tambah Occupation .....	42
Gambar 3.15 Rancangan Tampilan Ubah Occupation.....	43
Gambar 3.16 Rancangan Tampilan Item .....	44
Gambar 3.17 Rancangan Tampilan Tambah Item .....	44
Gambar 3.18 Rancangan Tampilan Ubah Item.....	45
Gambar 3.19 Rancangan Tampilan Data Rating.....	46
Gambar 3.20 Rancangan Tampilan Tambah Data Rating .....	46
Gambar 3.21 Rancangan Tampilan Rating Matrix .....	47
Gambar 3.22 Rancangan Tampilan Prior Dan Likelihood.....	48
Gambar 3.23 Rancangan Tampilan Classification Score Dan Reliability .....	49
Gambar 3.24 Rancangan Tampilan Rekomendasi Item.....	49

Gambar 4.1 Tabel Admin.....	50
Gambar 4.2 Tabel User .....	50
Gambar 4.3 Tabel Occupation .....	51
Gambar 4.4 Tabel Item .....	52
Gambar 4.5 Tabel Data .....	52
Gambar 4.6 Halaman Login.....	53
Gambar 4.7 Halaman Dashboard .....	54
Gambar 4.8 Halaman Admin .....	55
Gambar 4.9 Halaman User .....	55
Gambar 4.10 Halaman Occupation.....	56
Gambar 4.11 Halaman Item .....	56
Gambar 4.12 Halaman Data Rating .....	57
Gambar 4.13 Rating Matrix .....	57
Gambar 4.14 Prior Dan Likelihood.....	58
Gambar 4.15 Classification Score Dan Reliability .....	58
Gambar 4.16 Suggestion .....	59
Gambar 4.17 Listing Rating Matrix .....	66
Gambar 4.18 Listing Rating Matrix 2 .....	66
Gambar 4.19 Listing Prior Dan Likelihood .....	67
Gambar 4.20 Listing Data Demografis .....	68

## DAFTAR PERSAMAAN MATEMATIKA

PERSAMAAN (2.1) NAIVE BAYES .....	15
PERSAMAAN (2.2) MEAN SQUARED ERROR .....	18
PERSAMAAN (3.1) DESAIN METODE 1 .....	26
PERSAMAAN (3.2) DESAIN METODE 2 .....	26
PERSAMAAN (3.3) PENDEKATAN BERBASIS USER 1 .....	28
PERSAMAAN (3.4) PENDEKATAN BERBASIS USER 2 .....	28
PERSAMAAN (3.5) PENDEKATAN BERBASIS USER 3 .....	28
PERSAMAAN (3.6) PENDEKATAN BERBASIS ITEM 1 .....	28
PERSAMAAN (3.7) PENDEKATAN BERBASIS ITEM 2 .....	29
PERSAMAAN(3.8) PENDEKATAN BERBASIS ITEM 3 .....	29
PERSAMAAN (3.9) PREDIKSI .....	29
PERSAMAAN (3.10) RELIABILITAS .....	30





## INTISARI

Sistem rekomendasi sebagai salah satu alat yang banyak digunakan di beberapa situs e-commerce, *market place*, maupun *entertainment* kini telah banyak mengadopsi metode-metode dalam sistem pakar dan model-model matriks faktorisasi dalam disiplin ilmu matematika dalam segi pengembangannya, sehingga hasil yang didapat dalam proses pembuatan sistem rekomendasi itu sendiri dapat lebih optimal dengan mengurangi informasi yang berlebih bagi pengguna dalam web.

Secara garis besar ada dua pendekatan yang dapat dilakukan selama proses pengembangan sistem rekomendasi itu sendiri, yakni *memory-based* dimana hasil rekomendasinya tidak begitu akurat namun mudah dijelaskan, dan *model-based* dimana hasil rekomendasinya lebih akurat namun sulit untuk dipahami. Penulis mengajukan pemodelan *collaborative filtering* dengan menggunakan *naive bayes* pada algoritmanya dalam proses pengembangan sistem rekomendasi yang dapat merekomendasikan *item* dengan menggunakan keserupaan informasi *user* dan *item* yang ada berdasarkan data tertentu. Pada penelitian ini penulis menggunakan *dataset* pada MovieLens yang dikumpulkan oleh tim GroupLens Research Project dari Universitas Minnesota, Amerika Serikat. Data didapat dari laman resmi MovieLens dan dikumpulkan selama periode 7 bulan dari 19 September 1997 hingga 22 April 1998. Dan penulis menggunakan MSE (*Mean Squared Error*) untuk pengujiannya.

Adapun hasil yang didapat dari penelitian ini adalah bahwa metode pendekatan berbasis *item collaborative filtering* dapat diterapkan pada aplikasi sistem rekomendasi MovieLens, hasil yang didapat dari pengujian dinilai kurang akurat menggunakan MSE karena hasil prediksi masih diatas 1, setelah dihitung dengan jumlah *neighbor* yang berbeda prediksi *rating* dinilai cukup dengan rata-rata akurasi 53,8%, dan kurang akuratnya hasil dari penelitian setidaknya dikarenakan sedikitnya jumlah keragaman data (*sparsity*) yang digunakan sehingga sistem menampilkan hasil yang kurang akurat.

**Kata Kunci:** Sistem Rekomendasi, *Collaborative Filtering*, *Naive Bayes*, *Dataset*, *Mean Squared Error*

## ABSTRACT

*The recommendation system as one of the tools that is widely used in several e-commerce sites, market places, and entertainment now has many methods in expert systems and factorization matrix models in the discipline of mathematics in terms of its development, so that the results obtained in the process of making the recommendation system itself can be more optimal by reducing excess information for users on the web.*

*Broadly speaking, there are two approaches that can be taken during the development of the recommendation system itself, namely memory based where the results of the recommendations are not very accurate but easy, and based on a model where the results of the recommendations are more accurate but difficult to implement. The author proposes collaborative modeling using Naive Bayes in the algorithm in the process of developing a recommendation system that can assist items by using the similarity of user information and existing items based on certain data. In this study the authors used a dataset in MovieLens collected by the GroupLens Research Project team from the University of Minnesota, United States. Data obtained from the official website of MovieLens and collected for a period of 7 months from 19 September 1997 to 22 April 1998. And the author uses MSE (Mean Squared Error) for testing.*

*The results obtained from this study are that the item collaborative filtering-based approach method can be applied to the MovieLens recommendation system application, the results obtained from the test are considered inaccurate using MSE because the prediction results are still above 1, after being calculated by the number of neighbors that are different, the predicted rating is assessed enough with an average accuracy of 53,8%, and the inaccuracy of the results of the research is at least due to the small amount of data diversity (sparsity) used so that the system displays inaccurate results.*

**Keyword:** *Recommender System, Collaborative Filtering, Naive Bayes, Dataset, Mean Squared Error*