

**HYBRID REKOMENDASI SISTEM MENGGUNAKAN MATRIK
FAKTORISASI DAN LIGHTFM**

SKRIPSI



Disusun Oleh:

Eza Nanda

16.11.0855

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

HYBRID REKOMENDASI SISTEM MENGGUNAKAN MATRIK

FAKTORISASI DAN LIGHTFM

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika



Disusun Oleh:

Eza Nanda

16.11.0855

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

2020

PERSETUJUAN

SKRIPSI

HYBRID REKOMENDASI SISTEM MENGGUNAKAN MATRIK FAKTORISASI DAN LIGHTFM

yang dipersiapkan dan disusun oleh

Eza Nanda

16.11.0825

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 23 Oktober 2020

Dosen Pembimbing,

Hartatik, S.T., M.Cs

NIK. 190302232

PENGESAHAN

SKRIPSI

HYBRID REKOMENDASI SISTEM MENGGUNAKAN MATRIK FAKTORISASI DAN LIGHTFM

yang dipersiapkan dan disusun oleh

Eza Nanda

16.11.0855

telah dipertahankan di depan Dewan Penguji

pada tanggal 19 Oktober 2020

Susunan Dewan Penguji

Nama Penguji

Anna Baita, M.Kom

NIK. 190302256

Dina Maulina, M.Kom

NIK. 190302412

Hartatik, S.T, M.Cs

NIK. 190302232

Tanda Tangan

Skrripsi ini telah diterima sebagai salah satu persyaratan

untuk memperoleh gelar Sarjana Komputer

Tanggal 23 Oktober 2020

DEKAN FAKULTAS IMU KOMPUTER

Krisnawati, S.Si, M.T.

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI) dan ini 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 penulisan yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis dicantumkan 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, 27 Oktober 2020

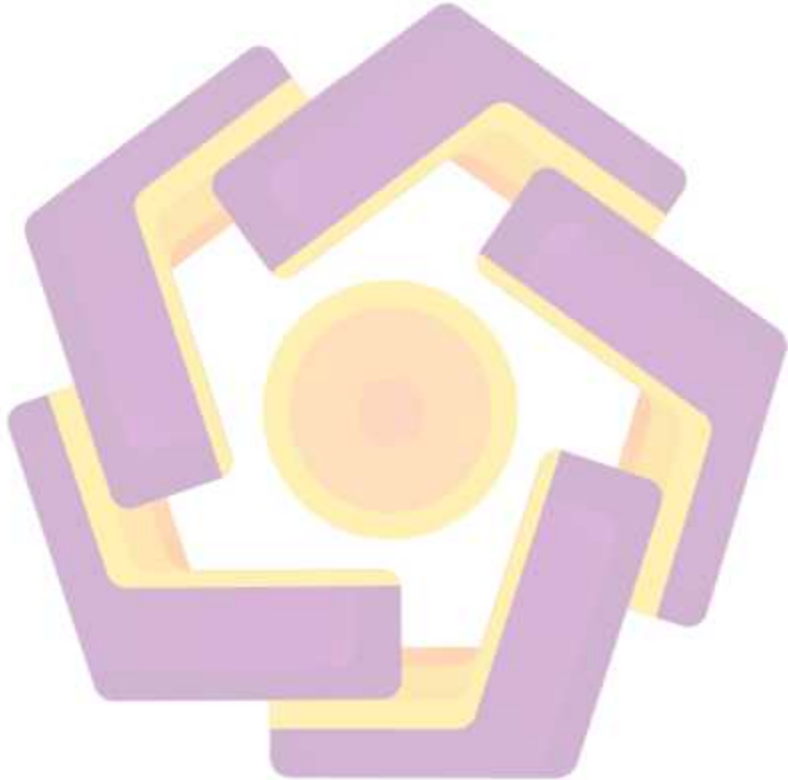


Eza Nanda

NIM 16.11.0855

MOTTO

Hidup hanya sekali, jadi nikmatilah apapun yang ada didepannya walau masalah selalu datang, tetap hadapi masalah tersebut sembari kita menikmati hidup ini. Hidup lah sebebas-bebasnya tanpa merugikan orang lain dan jangan lupa bersyukur.



PERSEMBAHAN

Alhamdulillah segala puji syukur atas kehadiran Allah SWT yang telah memberikan limpahan rahmat dan karunia-Nya sehingga penelitian ini dapat dilakukan dan diselesaikan dengan sebaik-baiknya. Saya juga ucapkan terimakasih untuk dukungan dan bantuan semua pihak yang membantu selesainya penelitian ini.

Skripsi ini saya persembahkan untuk kampus saya, Universitas AMIKOM Yogyakarta.

Untuk para pembaca, saya ucapkan terimakasih sebesar-besarnya. Semoga menjadi amal jariyah atas ilmu yang saya bagikan.

Untuk ibu Hartatik, yang sudah dengan sabar membimbing saya. Saya ucapkan terimakasih sedalam-dalamnya.

Dan untuk semua pihak yang telah banyak membantu dalam penyusunan skripsi ini yang tidak dapat saya sebutkan satu persatu

KATA PENGANTAR

Puji syukur penulis hanturkan atas kehadiran Allah SWT yang telah memberikan rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Hybrid Rekomendasi Sistem Menggunakan Matrik Faktorisasi dan LightFM” dengan sebaik-baiknya . Tidak lupa shalawat serta salam penulis haturkan kepada junjungan kita, nabi agung Nabi Muhammad SAW, yang telah membawa cahaya ilmu pengetahuan ke tengah-tengah kita.

Dengan selesainya skripsi ini, penulis mengucapkan banyak terimakasih kepada :

1. Bapak Prof. Dr. M. Suyanto selaku Rektor Universitas AMIKOM Yogyakarta
2. Bapak Sudarmawan, MT. selaku Dekan Fakultas Sains dan Teknologi, dan Ketua Program Studi SI Informatika
3. Ibu Hartatik, S.T., M.Cs. selaku pembimbing yang telah dengan sabra membimbing dan banyak memberikan kontribusi bagi penulis dalam pembuatan skripsi ini.
4. Para Dosen dan Staff Universitas AMIKOM Yogyakarta yang telah banyak memberi bantuan informasi, pengalaman dan kontribusi lainnya selama penulis berkuliah di kampus ini hingga terselesaikannya skripsi ini.
5. Orang tua terutama ibu saya dan sodara penulis yang telah memberikan dukungan terbaiknya selama kuliah.
6. Istikomah yang telah memberi semangat tanpa Lelah kepada saya sehingga dapat menyelesaikan skripsi ini.

7. Teman-teman kelas Informatika 13 Angkatan 2016 Universitas Amikom Yogyakarta yang bersama-sama menjadi pejuang gelar S.Kom.
8. Serta semua pihak lainnya yang tidak bisa dituliskan penulis satu per satu yang telah membantu selama pembuatan skripsi ini.

Demikian Skripsi ini, penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna dan masih banyak kekurangan. Hal tersebut dikarenakan keterbatasan ilmu dan pengetahuan yang dimiliki penulis semata. Oleh karena itu penulis mengharapkan kritik dan saran dari pembaca untuk menyempurnakan penulisan skripsi ini. Semoga skripsi ini dapat memberikan manfaat bagi pihak terkait dan pembaca pada umumnya. Akhir kata, semoga Allah SWT senantiasa melimpahkan rahmat serta hidayah-Nya kepada kita semua. Amin amin ya robbal 'alamin.

Yogyakarta 02 Oktober 2020

Eza Nanda

16.11.0855

DAFTAR ISI

PERSETUJUAN	iii
PENGESAHAN	iv
MOTTO	vi
PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABLE	xiii
DAFTAR ISTILAH	xiv
INTISARI	xv
ABSTRACT	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Penelitian	3
1.3 Batasan Penelitian	3
1.4 Maksud dan Tujuan Penelitian	4
1.5 Metode Penelitian	4
BAB II LANDASAN TEORI	7
2.1 Tinjauan Pustaka	7
2.2 Sistem Rekomendasi	10
2.2.1 Definisi Sistem Rekomendasi	10
2.2.3 Fungsi Sistem Rekomendasi	24
2.2.4 Content-Based Filtering	28
2.2.5 Collaborative Filtering	29
2.2.6 Hybrid Filtering	32
2.3 Matrix Factorization	36
2.4 LightFM	41
2.5 GroupLens	45
2.6 Kaggle	46
BAB III METODE PENELITIAN	47
3.1 Analisis Masalah	47
3.2 Analisis Kebutuhan Penelitian	48

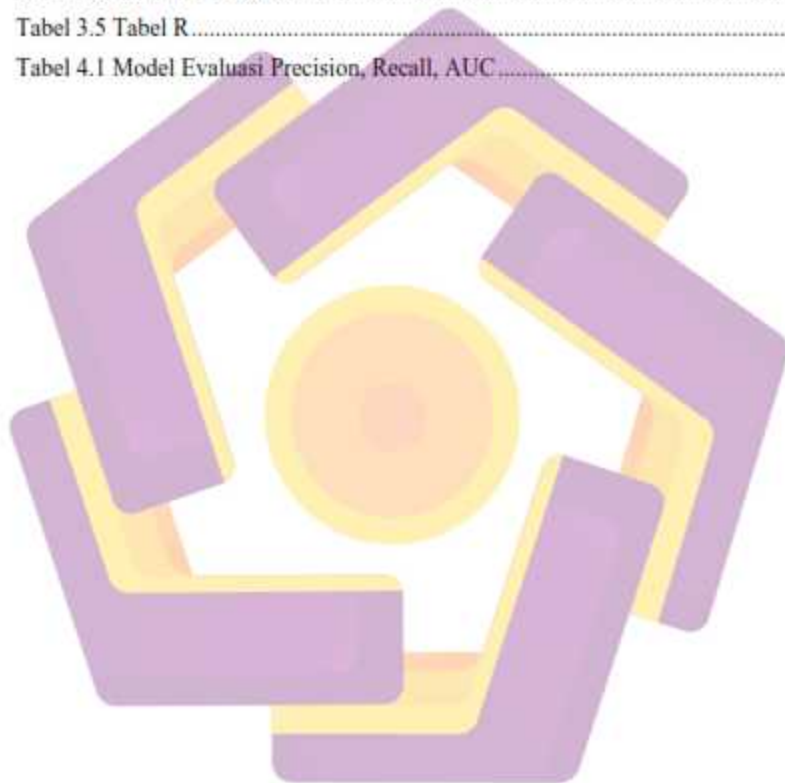
3.3	Arsitektur Umum	49
3.4	Dataset.....	52
3.5	Preprocessing.....	52
3.6	Model Matrix Factorization	53
3.7	Analisis Algoritma Rank Matrix Factorization.....	53
3.8	Analisis Hasil.....	64
BAB IV IMPLEMTASI DAN PENGUJIAN		65
4.1	Implementasi Sistem.....	65
4.2	Implementasi Data.....	65
4.3	Pre-processing Data.....	67
4.4	Penentuan nilai Epoch.....	68
4.5	Membuat matrik faktorisasi.....	69
4.6	Hasil Rekomendasi.....	69
4.7	Evaluasi Model.....	75
4.7.1	Pengaruh Jumlah Epoch.....	76
4.7.2	Precision @ k.....	77
4.7.3	Recall @ k.....	79
4.7.4	Diagram AUC-ROC.....	81
4.7.5	Ringkasan Evaluasi Model.....	82
BAB V PENUTUP		84
5.1	Kesimpulan	84
5.2	Saran.....	85
DAFTAR PUSTAKA.....		87

DAFTAR GAMBAR

Gambar 2.1 Teknik Rekomendasi.....	21
Gambar 2.2 Teknik Collaborative Filtering.....	29
Gambar 2.3. Alur Proses Hybrid Liner Combination.....	33
Gambar 2.4. Alur Proses Hybrid Sequential Combination.....	33
Gambar 2.5 Alur Proses Item-based Clustering Hybrid Method (ICHM).....	34
Gambar 2.6 Contoh Perhitungan Matrix Factorization.....	37
Gambar 3.1 Arsitektur umum.....	51
Gambar 4.1 Ratings dataset.....	66
Gambar 4.2 Movie dataset.....	66
Gambar 4.3 proses preprocessing data.....	68
Gambar 4.4 Perbandingan nilai epoch.....	68
Gambar 4.5 Script untuk model matrik faktorisasi.....	69
Gambar 4.6 <i>Item</i> Recommendation to user dataset movielens.....	70
Gambar 4.7 Item Recommendation to <i>user</i> dataset Kaggle.....	72
Gambar 4.8 User rekomendasi to item.....	73
Gambar 4.9 Item recommendation to item.....	74
Gambar 4.10 Akurasi AUC Berdasarkan Adagrad.....	76
Gambar 4.11 Akurasi AUC Berdasarkan Adadelata.....	77
Gambar 4.12 Akurasi Precision @ 4 berdasarkan Loss Function.....	78
Gambar 4.13 Perbandingan jumlah k pada precision.....	79
Gambar 4.14 Akurasi Recall @ 4 berdasarkan Loss Function.....	80
Gambar 4.15 Perbandingan jumlah k pada recall.....	81
Gambar 4.16 Akurasi AUC-ROC berdasarkan Loss Function.....	82

DAFTAR TABLE

Tabel 3.1 Tabel R.....	55
Tabel 3.2 Proses Dekomposisi.....	56
Tabel 3.3 Prediksi Hasil Nilai R.....	58
Tabel 3.4 Contoh Prediksi Matrik.....	59
Tabel 3.5 Tabel R.....	60
Tabel 4.1 Model Evaluasi Precision, Recall, AUC.....	83



DAFTAR ISTILAH

- Cold-start** : Kondisi dimana sistem tidak dapat menarik kesimpulan untuk pengguna atau item yang baru yang belum memberikan informasi (rate) yang memadai
- Dekomposisi** : Proses perubahan menjadi bentuk lebih sederhana
- Faktorisasi** : Proses dekomposisi menjadi perkalian matriks menjadi matrik baru yang ketika dikalikan bersama menghasilkan bilangan aslinya
- Epoch** : Banyaknya proses yang dilalui dataset dalam satu kali proses yang dilewati satu kali proses forward dan backward melewati semua node neural network. Karena setiap satu epoch terlalu besar untuk memasukan semua data kedalam komputer sekaligus, maka dibaginya jadi beberapa batch yang lebih kecil.
- Learning rate** : Parameter dari gradient descent
- Gradient descent** : Algoritma Untuk mengoptimalkan iterasi yang digunakan pada Machine Learning untuk menemukan hasil yang terbaik
- Sparsity** : Kondisi tidak memiliki cukup sesuatu atau jarangny data dalam sebuah model data
- Skalabilitas** : Kemampuan sistem, atau proses untuk menangani penambahan beban yang diberikan, atau potensinya untuk ditingkatkan guna menangani penambahan beban tersebut

INTISARI

Perkembangan terus terjadi pada dunia teknologi, yang berdampak besar pada bidang komunikasi dan informasi. Hal ini dibuktikan dengan terus meningkatnya penggunaan internet pada kehidupan di masyarakat. Selain internet dunia perfilman juga mengalami perkembangan pesat. Peningkatan ini diikuti oleh meningkatnya minat menonton film oleh masyarakat. Namun dengan berkembangnya dunia perfilman, jumlah film yang beredar pun semakin banyak, sehingga masyarakat memerlukan informasi yang dapat membantu dalam menentukan film yang diinginkan. Oleh karena itu, dibutuhkan sebuah sistem yang dapat memberikan rekomendasi untuk menghemat waktu dalam menentukan film.

Pada penelitian ini, sistem rekomendasi yang dibangun menggunakan metode hybrid rekomendasi menggunakan algoritma matrik faktorisasi dengan bantuan LightFM. Pengujian dilakukan dengan menggunakan dataset publik kemudian dilakukan proses *preprocessing*, kemudian dilakukan penentuan jumlah *epoch*, yang nantinya akan digunakan dalam menentukan model matrik faktorisasi.

Kesimpulan yang didapat dari penelitian ini adalah akurasi dari model yang telah dievaluasi menggunakan *AUC-ROC score*, *Precision* dan *Recall*. Menghasilkan akurasi *AUC-ROC score* dengan *loss function* WARP dan kombinasi Adagrad dengan *score* sebesar 93,6% Sedangkan untuk *precision @ 4* cukup rendah dengan nilai 33,2% dan *Recall @ 4* dengan nilai 12,1% untuk nilai $k=4$.

Kata Kunci : Rekomendasi sistem, Matrik faktorisasi, LightFM, Hybrid Rekomendasi, Film

ABSTRACT

Development continues to occur in the world of technology, which has a major impact on the field of communication and information. This is evidenced by the increasing use of the internet in people's lives. Apart from the internet, the world of cinema is also experiencing rapid development. This increase was followed by increased interest in watching films by the public. However, with the development of the world of film, the number of films that appear is increasing, so that people need information that can help in determining the desired film. Therefore, we need a system that can provide recommendations to save time in determining films.

In this study, the recommendation system was built using a hybrid recommendation method using a factorization matrix algorithm with the help of LightFM. The test was carried out using a public dataset, then a preprocessing process was carried out, then the number of epochs was determined, which would later be used in determining the factorization matrix model.

The conclusion obtained from this study is the accuracy of the model that has been evaluated using the AUC-ROC score, Precision and Recall. The results are AUC-ROC score accuracy with the WARP loss function and Adagrad combination with a score of 93.6% while for precision @ 4 is quite low, with a value of 33.2% and Recall @ 4 with a value of 12.1% for the value of $k = 4$.

Keywords: Recommendation system, Factorization matrix, LightFM, Hybrid Recommendation, Film