

OPTIMISASI COLLABORATIVE FILTERING BERBASIS HYBRID MENGUNAKAN MATRIX FACTORIZATION, FEEDFORWARD NEURAL NETWORK DAN XGBOOST

TESIS

Oleh:

**FILIMANTAPTIUS GULO
NIM. 231232211**



**PROGRAM STUDI S-2 TEKNOLOGI INFORMASI
FAKULTAS INFORMATIKA
UNIVERSITAS MIKROSKIL
MEDAN
2025**

OPTIMIZATION OF HYBRID-BASED COLLABORATIVE FILTERING USING MATRIX FACTORIZATION, FEEDFORWARD NEURAL NETWORK AND XGBOOST

THESIS

By:

**FILIMANTAPTIUS GULO
ID NUMBER. 231232211**



**MAJOR OF S-2 INFORMATION TECHNOLOGY
FACULTY OF INFORMATICS
UNIVERSITAS MIKROSKIL
MEDAN
2025**

LEMBARAN PENGESAHAN

**OPTIMISASI COLLABORATIVE FILTERING BERBASIS HYBRID
MENGUNAKAN MATRIX FACTORIZATION, FEEDFORWARD
NEURAL NETWORK DAN XGBOOST**

TESIS

Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Magister
Program Studi S-2 Teknologi Informasi

Oleh:

FILIMANTAPTIUS GULO
NIM. 231232211

Disetujui Oleh:

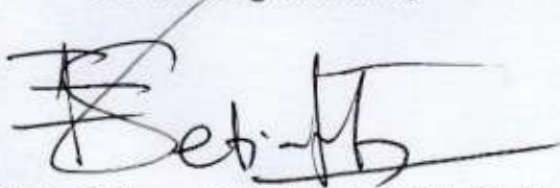
Dosen Pembimbing,



Dr. Ronsen Purba, M.Sc.

Medan, 3 September 2025
Diketahui dan Disahkan Oleh:

Ketua Program Studi
S-2 Teknologi Informasi,



Ir. Erwin Setiawan Panjaitan, M.M.S.I., Ph.D.

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-2 Teknologi Informasi Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : FILIMANTAPIUS GULO

NIM : 231232211

Saya telah melaksanakan penelitian dan penulisan Tesis dengan judul dan tempat penelitian sebagai berikut:

Judul Tesis : Optimisasi Collaborative Filtering berbasis Hybrid menggunakan Matrix Factorization, Feedforward Neural Network dan XGBoost.

Tempat Penelitian : -

Alamat Tempat Penelitian : -

No. Telp. Tempat Penelitian : -

Sehubungan dengan Tesis tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Tesis tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Tesis saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tesis saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Tesis saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 25 Juni 2025

Saya yang membuat pernyataan,



Filimantaptius Gulo

OPTIMISASI COLLABORATIVE FILTERING BERBASIS HYBRID MENGGUNAKAN MATRIX FACTORIZATION, FEEDFORWARD NEURAL NETWORK DAN XGBOOST

Abstrak

Rekomendasi *collaborative filtering* telah digunakan secara luas dalam berbagai aplikasi digital, namun masih menghadapi tantangan seperti *cold-start*, *first-rater* dan kurangnya akurasi rekomendasi akibat ketidakmampuannya dalam menangkap hubungan kompleks antara *user* dan item. Penelitian ini bertujuan untuk mengoptimisasi metode *collaborative filtering* berbasis *hybrid* melalui integrasi tiga pendekatan utama, yaitu *Matrix Factorization*, *Feedforward Neural Network* berbasis *Multi-Layer Perceptron*, dan *XGBoost*. Metodologi yang digunakan mencakup tahapan membaca dataset, *preparation*, *tuning hyperparameter*, hingga pemodelan dan evaluasi menggunakan metrik *MAE*, *RMSE*, dan *R² Score*. Penelitian dilakukan pada dua *dataset* berbeda, yakni film dan hotel, untuk menguji generalisasi model. Hasil penelitian menunjukkan bahwa model *hybrid* yang diusulkan mampu mengatasi keterbatasan *collaborative filtering non-hybrid* dan secara signifikan meningkatkan akurasi rekomendasi. Temuan ini memberikan kontribusi terhadap pengembangan sistem rekomendasi yang lebih adaptif dan personal, serta membuka peluang untuk eksplorasi lebih lanjut dalam integrasi metode *machine learning* pada sistem rekomendasi di era data besar.

Kata Kunci: *collaborative filtering, matrix factorization, feedforward neural network, xgboost*

Abstract

Collaborative filtering recommendations have been widely used in various digital applications, but still face challenges such as cold-start, first-rater issues, and a lack of recommendation accuracy due to its inability to capture the complex relationships between users and items. This study aims to optimize the collaborative filtering method through a hybrid approach by integrating three main techniques: Matrix Factorization, Feedforward Neural Network based on Multi-Layer Perceptron, and XGBoost. The methodology includes stages such as dataset reading, data preparation, hyperparameter tuning, modeling, and evaluation using metrics such as MAE, RMSE, and R² Score. The study was conducted across two different datasets—movies and hotels—to assess the model's generalization capabilities. The results indicate that the proposed hybrid model successfully addresses the limitations of non-hybrid collaborative filtering and significantly improves the accuracy of recommendations. These findings contribute to the development of more adaptive and personalized recommendation systems and open new opportunities for further exploration of machine learning integration in recommendation systems in the era of big data.

Keywords: *collaborative filtering, matrix factorization, feedforward neural network, xgboost*

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, penulis dapat menyelesaikan tesis yang berjudul **“Optimisasi Collaborative Filtering berbasis Hybrid menggunakan Matrix Factorization, Feedforward Neural Network dan XGBoost”** dengan baik.

Tesis ini disusun sebagai salah satu syarat untuk menyelesaikan Program Studi Magister Teknologi Informasi di Universitas Mikroskil Medan.

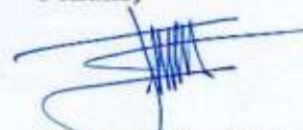
Penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada semua pihak yang telah memberikan dukungan, baik secara langsung maupun tidak langsung, dalam proses penyusunan tesis ini. Secara khusus, penulis mengucapkan terima kasih kepada:

1. Bapak Dr. Ronsen Purba, M.Sc., selaku Dosen Pembimbing yang telah memberikan bimbingan dan arahan kepada penulis dalam penyelesaian tesis ini.
2. Bapak Muhammad Fermi Pasha, B.Sc., M.Sc., Ph.D., selaku Dosen Pendamping Pembimbing yang telah turut membimbing dan mendukung penulis dalam proses penyelesaian tesis ini.
3. Bapak Hardy, S.Kom., M.Sc., Ph.D., selaku Rektor Universitas Mikroskil Medan.
4. Bapak Sunaryo Winardi, S.Kom., M.T., selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
5. Bapak Ir. Erwin Setiawan Panjaitan, M.M.S.I., Ph.D., selaku Ketua Program Studi S-2 Teknologi Informasi Universitas Mikroskil Medan.
6. Bapak atau Ibu Dosen Universitas Mikroskil Medan yang telah membantu proses penulisan tesis ini.
7. Anggota keluarga, teman, saudara dan semua pihak yang terus memberikan dukungan penuh kepada penulis selama proses penulisan tesis ini.

Penulis menyadari bahwa masih terdapat berbagai kekurangan dalam penulisan tesis ini. Oleh karena itu, kritik dan saran yang bersifat membangun sangat diterima. Akhir kata, penulis berharap tesis ini dapat memberikan manfaat bagi masyarakat. Terima kasih.

Medan, 25 Juni 2025

Penulis,



Filimantaptius Gulo

DAFTAR ISI

ABSTRAK.....	i
KATA PENGANTAR	ii
DAFTAR ISI	iii
DAFTAR GAMBAR	v
DAFTAR TABEL	vi
DAFTAR LAMPIRAN	vii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan	3
1.4 Manfaat	4
1.5 Ruang Lingkup	4
BAB II KAJIAN LITERATUR.....	5
2.1 Tinjauan Pustaka.....	5
2.1.1 Rekomendasi Berbasis Hybrid.....	5
2.1.2 Collaborative Filtering	6
2.1.3 Matrix Factorization.....	8
2.1.4 Feedforward Neural Network.....	9
2.1.5 Handcrafted Features.....	11
2.1.6 XGBoost.....	12
2.2 Penelitian Terdahulu	13
2.3 Kerangka Konseptual.....	20
BAB III METODOLOGI PENELITIAN.....	22
3.1 Analisis Masalah	22
3.2 Metode Penelitian	23
3.2.1 Membaca Dataset	24
3.2.2 Splitting Data.....	24
3.2.3 Data Preparation	24
3.2.4 Tuning Hyperparameter	24
3.2.5 Pemodelan	25
3.2.6 Pengujian.....	29
3.3 Data yang Digunakan.....	30

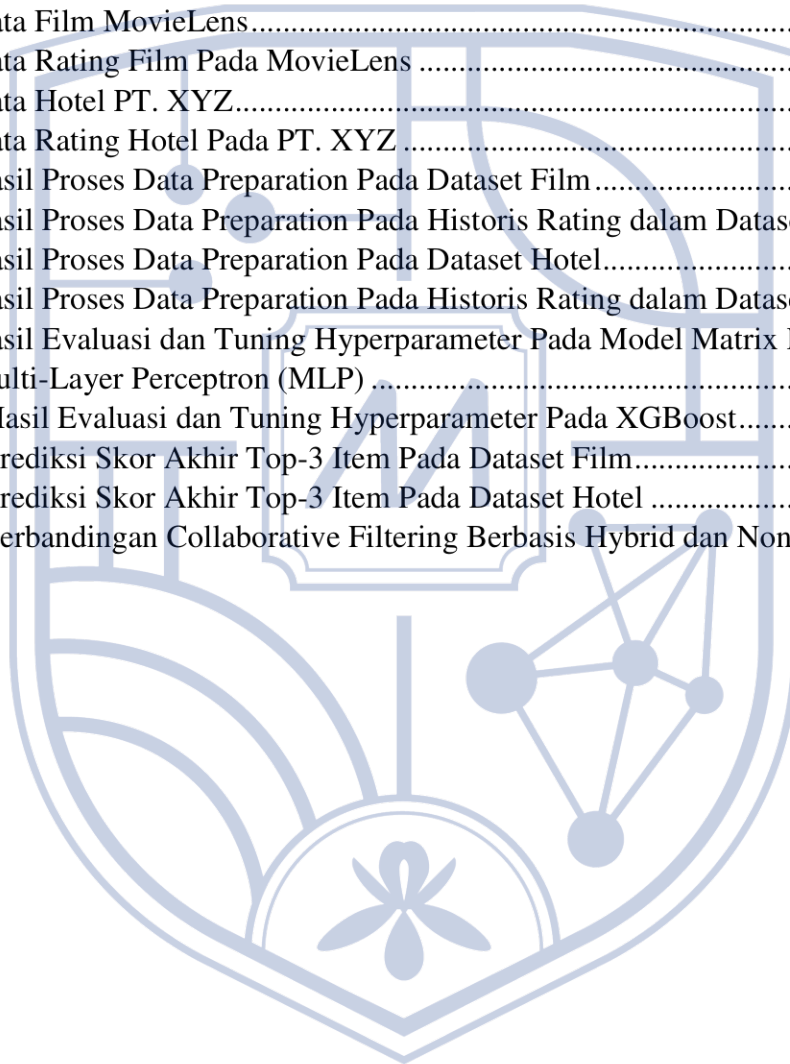
3.4 Alat-alat Penelitian	32
3.5 Teknik Analisis Data	32
BAB IV HASIL DAN PEMBAHASAN.....	34
4.1 Hasil.....	34
4.1.1 Tahapan Awal Membaca Dataset.....	34
4.1.2 Tahap Splitting Data.....	38
4.1.3 Proses Data Preparation	39
4.1.4 Variabel Tuning Hyperparameter.....	42
4.1.5 Pemodelan Collaborative Filtering Berbasis Hybrid	43
4.1.6 Hasil Pengujian	52
4.2 Pembahasan	56
4.2.1 Analisis Terhadap Masalah Cold-Start dan First-Rater	56
4.2.2 Perbandingan Kinerja Rekomendasi Collaborative Filtering Hybrid dan Non-Hybrid.....	57
4.2.3 Analisis Komparatif Akurasi Model Terhadap Studi Sebelumnya.....	58
4.2.4 Generalitas Model pada Berbagai Domain	58
BAB V PENUTUP.....	59
5.1 Kesimpulan	59
5.2 Saran	59
DAFTAR PUSTAKA	60
DAFTAR RIWAYAT HIDUP.....	63

DAFTAR GAMBAR

Gambar 2.1 Diagram User-based [6].....	7
Gambar 2.2 Diagram Item-based [6].....	8
Gambar 2.3 Systematic Literature Review.....	14
Gambar 2.4 Kerangka Konseptual.....	20
Gambar 3.1 Metode Penelitian	23
Gambar 3.2 Model Hybrid Collaborative Filtering.....	25
Gambar 3.3 Data Film dan Hotel Beserta Rating Review.....	26
Gambar 4.1 Kode Program Untuk Membaca Dataset Film.....	34
Gambar 4.2 Kode Program Untuk Membaca Dataset Hotel	34
Gambar 4.3 Kode Program Untuk Splitting Data.....	38
Gambar 4.4 Kode Program Untuk Data Preparation.....	39
Gambar 4.5 Kode Program Tuning Hyperparameter Pada Matrix Factorization.....	42
Gambar 4.6 Kode Program Tuning Hyperparameter Pada Multi-Layer Perceptron.....	42
Gambar 4.7 Kode Program Tuning Hyperparameter Pada XGBoost	43
Gambar 4.8 Kode Program Matrix Factorization.....	44
Gambar 4.9 Kode Program Persiapan Input Pada Multi-Layer Perceptron	45
Gambar 4.10 Kode Program Pembangunan Arsitektur Multi-Layer Perceptron	46
Gambar 4.11 Kode Program Pelatihan Multi-Layer Perceptron	47
Gambar 4.12 Kode Program Handcrafted Features.....	51
Gambar 4.13 Hasil Handcrafted Features.....	52
Gambar 4.14 Kode Program XGBoost.....	52

DAFTAR TABEL

Tabel 2.1 Ringkasan dari tujuh model berbasis hybrid	5
Tabel 2.2 Rangkuman Penelitian Terdahulu	17
Tabel 3.1 Struktur Data Film dan Hotel	31
Tabel 3.2 Struktur Data Historis Rating Pada Film dan Hotel	31
Tabel 4.1 Data Film MovieLens	35
Tabel 4.2 Data Rating Film Pada MovieLens	36
Tabel 4.3 Data Hotel PT. XYZ	37
Tabel 4.4 Data Rating Hotel Pada PT. XYZ	37
Tabel 4.5 Hasil Proses Data Preparation Pada Dataset Film	40
Tabel 4.6 Hasil Proses Data Preparation Pada Historis Rating dalam Dataset Film	40
Tabel 4.7 Hasil Proses Data Preparation Pada Dataset Hotel	41
Tabel 4.8 Hasil Proses Data Preparation Pada Historis Rating dalam Dataset Hotel	41
Tabel 4.9 Hasil Evaluasi dan Tuning Hyperparameter Pada Model Matrix Factorization (MF) dan Multi-Layer Perceptron (MLP)	47
Tabel 4.10 Hasil Evaluasi dan Tuning Hyperparameter Pada XGBoost	53
Tabel 4.11 Prediksi Skor Akhir Top-3 Item Pada Dataset Film	55
Tabel 4.12 Prediksi Skor Akhir Top-3 Item Pada Dataset Hotel	56
Tabel 4.13 Perbandingan Collaborative Filtering Berbasis Hybrid dan Non-hybrid	57



DAFTAR LAMPIRAN

Lampiran 1 Gambar Diagram Tebar (Scatter Plot) Perbandingan Nilai Rating Aktual dan Prediksi XGBoost	62
--	----

