



UNDERGRADUATE THESIS

A MOVIE RECOMMENDATION SYSTEM BASED ON USER RATING SIMILARITY USING MATRIX FACTORIZATION

SATYA AGNI PREMA RASWIDHYANTORO
NPM 19081010109

THESIS ADVISORS

Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T.
Andreas Nugroho Sihananto, S.Kom., M.Kom.

**MINISTRY OF HIGHER EDUCATION, SCIENCE, AND TECHNOLOGY
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FACULTY OF COMPUTER SCIENCE
INFORMATICS STUDY PROGRAM
SURABAYA
2026**

APPROVAL SHEET

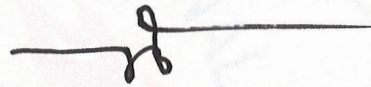
A MOVIE RECOMMENDATION SYSTEM BASED ON USER RATING SIMILARITY USING MATRIX FACTORIZATION

By:
SATYA AGNI PREMA RASWIDHYANTORO
NPM. 19081010109

Has been defended before, and accepted by, the Board of Assessors of the Thesis Examination of the Informatics Study Program, Faculty of Computer Science, Universitas Pembangunan Nasional Veteran Jawa Timur, on May 18th, 2026:

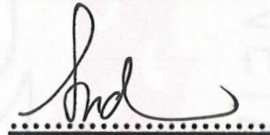
Approved,

Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T
NIP. 19830310 202121 1 006



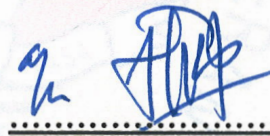
(Advisor I)

Andreas Nugroho Sihananto, S.Kom., M.Kom.
NIP. 19900412 202406 1 003



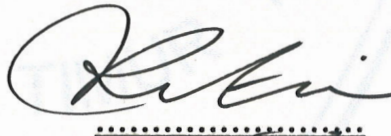
(Advisor II)

Henni Endah Wahanani, ST. M.Kom.
NIP. 19780922 202121 2 005



(Head Assessor)

Retno Mumpuni, S.Kom., M.Sc
NIP. 198707162025212045



(Assessor I)

Acknowledged by,
Dean of the Faculty of Computer Science



Prof. Dr. Ir. Novirina Hendrasarie, MT.
NIP. 19681126 199403 2 001

APPROVAL SHEET

A MOVIE RECOMMENDATION SYSTEM BASED ON USER RATING SIMILARITY USING MATRIX FACTORIZATION

By :
SATYA AGNI PREMA RASWIDHYANTORO
NPM. 19081010109

Approved to proceed to the Thesis Examination

Approved by,

**Coordinator of Informatics Study Program
Faculty of Computer Science**



Dr. Intan Yuniar Purbasari, S.Kom. MSc.
NIP. 19800602 202521 2 029

STATEMENT OF ORIGINALITY

I am the undersigned:

Student Name : Satya Agni Prema Raswidhyantoro
NPM : 19081010109
Degree Program : Bachelor (S1)
Study Program : Informatics
Faculty : Faculty of Computer Science

Hereby declares that this undergraduate thesis contains no part of any other scientific work that has been submitted to obtain an academic degree at any higher education institution. Furthermore, it does not contain any work or opinions previously written or published by others, except for those which are explicitly cited in this thesis and listed completely in references.

And I declare that this scientific document is free from elements of plagiarism. If in the future indications of plagiarism are found in this Thesis, I am willing to accept sanctions in accordance with the applicable laws and regulations.

Thus, I made this statement without any coercion from anyone and to be used as it should.



Surabaya, May 18th, 2026
Declarant,



Satya Agni Prema R.
NPM. 19081010109

ABSTRACT

Student Name / NPM : Satya Agni Prema Raswidhyantoro / 19081010109
Thesis Title : A Movie Recommendation System Based on User
Rating Similarity Using Matrix Factorization
Supervisor : 1. Dr. Ir. Mohammad Idhom, SP, S.Kom., MT
2. Andreas Nugroho Sihananto, S.Kom., M.Kom

The growth of streaming platforms and digital entertainment services has increased the number of available movies, making it difficult for users to identify movies that match their preferences. This research aims to develop a movie recommendation system based on user rating similarity using the Matrix Factorization algorithm. The system is designed to allow users to input a desired rating and select a movie genre, after which the system generates a Top-N recommendation list according to those preferences.

The dataset used in this research is MovieLens, consisting of 100,836 rating records, 610 unique users, and 9,724 unique movies. The dataset has a sparsity level of 98.30%, making Matrix Factorization suitable for learning latent patterns between users and movies from an incomplete rating matrix. The model was trained using training data and monitored using validation data with an early stopping mechanism. The training process achieved a Best Val MSE of 0.7243 and stopped at the 19th epoch. Evaluation on the test data produced an RMSE of 0.8657 and an MAE of 0.6617, indicating that the average rating prediction error is below one point on the MovieLens rating scale.

The recommendation mechanism uses the input rating to find historical users whose average ratings are close to the user's preference within a selected genre. Genre information is used as a movie candidate filter to ensure that the recommendations remain aligned with the selected category. In addition, the application provides an advanced exploration feature through clickable recommendation cards, allowing the selected movie to serve as an additional reference for the next recommendation process. The system is implemented as a Flask-based web application that displays recommendations in movie cards containing the title, genre, predicted rating, recommendation score, poster or placeholder, and score visualization. Based on the implementation and testing results, the system is able to generate movie recommendations that match the user's rating and genre inputs while presenting the results in an informative and interactive interface.

Keywords: Recommender System, Matrix Factorization, Collaborative Filtering, Content-Based Filtering, Hybrid Recommendation, MovieLens, Flask.

ACKNOWLEDGEMENTS

The author expresses praise and gratitude to Almighty God for His grace, guidance, and blessings, which enabled the successful completion of this undergraduate thesis entitled **“A Movie Recommendation System Based on User Rating Similarity Using Matrix Factorization.”**

The author would like to express sincere gratitude to Dr. Ir. Mohammad Idhom, S.P., S.Kom., M.T., as the principal thesis advisor, for the time, guidance, advice, and motivation provided throughout the preparation of this thesis. The author also received valuable moral, spiritual, and material support from various parties. Therefore, the author would like to express sincere appreciation to :

1. Prof. Dr. Ir. Novirina Hendrasarie, M.T., Dean of the Faculty of Computer Science, Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Dr. Intan Yuniar Purbasari, S.Kom., M.Sc., Head of the Informatics Study Program, Faculty of Computer Science, Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Andreas Nugroho Sihananto, S.Kom., M.Kom., as the second thesis advisor.
4. Henni Endah Wahanani, S.T., M.Kom., and Retno Mumpuni, S.Kom., M.Sc., as members of the thesis examination committee.
5. The author’s parents, Mr. Tegoeh and Mrs. Sista, for their continuous support, encouragement, and prayers.

The author acknowledges that this undergraduate thesis still has limitations and shortcomings. Therefore, constructive criticism and suggestions from all parties are highly appreciated for its improvement. Despite these limitations, the author sincerely hopes that this thesis will be beneficial to readers in general and to the author in particular.

Surabaya, May 18th 2026

Author

Table Of Contents

APPROVAL SHEETS.....	iii
STATEMENT OF ORIGINALITY	v
ABSTRACT	vii
ACKNOWLEDGEMENTS.....	xi
TABLE OF CONTENTS.....	xiii
LIST OF FIGURES	xiv
LIST OF TABLES	xviii
LIST OF SYMBOLS.....	xix
CHAPTER I INTRODUCTION.....	1
1.1. Background	1
1.2. Research Questions	3
1.3. Research Objectives	4
1.4. Research Benefits	4
1.5. Scope and Limitations	5
CHAPTER II LITERATURE REVIEW	6
2.1. Previous Studies	6
2.2. Theoretical Background	10
2.2.1. Recommender Systems	10
2.2.2. Collaborative Filtering	10
2.2.3. Content Based Filtering.....	11
2.2.4. Hybrid Recommendation	11
2.2.5. Data Sparsity	12
2.2.6. Cold Start.....	13
2.2.7. Matrix Factorization	13
2.2.8. Cosine Similarity.....	14
2.2.9. MovieLens Dataset.....	15
2.2.10. Recommender System Evaluation	15
2.2.11. Rating Prediction and Top-N Recommendation	15
2.2.12. Genre as Movie Metadata	16
2.2.13. Hybrid Approaches in Movie Recommendation Systems	17

2.2.14. User Cold Start in Recommendation Applications	18
2.2.15. Supporting Research Tools	19
CHAPTER III SYSTEM DESIGN AND IMPLEMENTATION.....	21
3.1. Research Methodology.....	21
3.2. Research Dataset	22
3.3. Data Preprocessing	23
3.4. Data Splitting.....	25
3.5. Matrix Factorization Model Design	25
3.6. Training Process and Early Stopping.....	26
3.7. User Profile Construction Based on Rating and Genre.....	28
3.8. Rating- and Genre-Based Recommendation Mechanism	29
3.8.1. Pseudocode for the Rating- and Genre-Based Recommendation Mechanism.....	31
3.9. Advanced Exploration Through Recommendation-Card Clicks	32
3.9.1. Pseudocode for the Recommendation-Card Click Feature	33
3.10. Recommendation Calculation Example	33
3.11. Movie Poster Integration and Caching Strategy	35
3.12. Flask-Based Web Application Implementation	36
3.12.1. User Interface Design.....	39
3.12.2. Application Runtime Flow	40
3.13. System Development Environment.....	41
3.14. Model Evaluation	41
3.15. System Functional Testing	42
CHAPTER IV RESULTS AND DISCUSSION	45
4.1. Overview of Testing and Achieved Results.....	44
4.2. Dataset Processing Results and Data Characteristics.....	45
4.2.1. Rating Distribution in the Dataset.....	47
4.2.2. Distribution of Movies by Genre	48
4.3. Training, Validation, and Test Data Splitting	50
4.4. Results of User Profile Construction Based on Rating and Genre.....	51
4.5. Matrix Factorization Model Training Results.....	52

4.6.	Model Evaluation on the Test Set Using RMSE and MAE	54
4.6.1.	Discussion of RMSE and MAE Results.....	56
4.7.	Recommendation Results Based on Rating and Genre.....	56
4.7.1.	Analysis of Recommendations for a Rating of 3.5 and the Drama Genre	59
4.7.2.	Analysis of Recommendations for a Rating of 4.0 and the Thriller Genre	61
4.8.	Comparison of Recommendation Results Across Different Genres.....	62
4.9.	Results of the Advanced Exploration Feature Through Recommendation- Card Clicks	65
4.9.1.	Comparison of Recommendations Before and After a Card IsClicked	66
4.10.	Results of Poster, Placeholder, Genre Badge, and Visual Score Integration.....	68
4.11.	Functional Testing Results	71
4.12.	Discussion of System Limitations.....	76
CHAPTER V CONCLUSIONS AND RECOMMENDATIONS.....		78
5.1.	Conclusions	78
5.2.	Recommendations	80
REFERENCES.....		82
ATTACHMENTS		84

LIST OF FIGURES

Figures 2.1	Illustration of Rating Matrix Decomposition in Matrix Factorization.....	14
Figures 3.1	Research Methodology Flowchart	21
Figures 3.2	Data Preprocessing Flowchart for Ratings and Movie Metadata.....	24
Figures 3.3	Matrix Factorization Architecture.....	25
Figures 3.4	Training MSE and Validation MSE per Epoch with the Early-Stopping Point.....	26
Figures 3.5	User Profile Construction Based on Rating and Genre.....	27
Figures 3.6	Rating- and Genre-Based Recommendation Flowchart....	29
Figures 3.7	Recommendation-Card Click Workflow as an Exploration Reference.....	31
Figures 3.8	Illustration of Recommendation Score Calculation.....	36
Figures 3.9	Movie Poster Integration and Caching Workflow.....	37
Figures 3.10	Rating, Genre, and Top-N Input Page.....	38
Figures 3.11	Recommendation Results Displayed as Movie Cards.....	39
Figures 3.12	Exploration Mode after a Movie Card Is Clicked.....	39
Figures 3.13	RMSE and MAE Evaluation Results.....	43
Figures 4.1	Main Application Page for Rating, Genre, and Top-N Input.....	45
Figures 4.2	Dataset Statistics after Preprocessing.....	47
Figures 4.3	Rating Distribution in the Dataset.....	48
Figures 4.4	Number of Movies by Genre.....	50
Figures 4.5	Recommendation Profile Summary in the Application	51
Figures 4.6	Model Training Log Output in Google Colab.....	53
Figures 4.7	Training MSE and Validation MSE	54

Figures 4.8	RMSE and MAE Results.....	55
Figures 4.9	Recommendation Results for a Rating of 4.0 and the Comedy Genre	57
Figures 4.10	Recommendation Results for a Rating of 3.5 and the Drama Genre	59
Figures 4.11	Recommendation Results for a Rating of 4.0 and the Thriller Genre	60
Figures 4.12	Recommendation Results for a Rating of 4.5 and the Action Genre.....	62
Figures 4.13	Top-5 Test Results.....	64
Figures 4.14	Top-20 Test Results.....	65
Figures 4.15	Results after Duck Soup (1933) Is Clicked as the Exploration Reference.....	66
Figures 4.16	Recommendation Results before a Movie Card Is Clicked..	67
Figures 4.17	Recommendation Results after a Movie Card Is Clicked.....	68
Figures 4.18	Recommendation Cards with Genre Badges and Visual Scores	68
Figures 4.19	Placeholder Displayed When a Movie Poster Is Unavailable.....	69
Figures 4.20	Evidence of Functional Test F-01 for Valid Rating and Genre Input	71
Figures 4.21	Evidence of the Top-N Functional Test.....	71
Figures 4.22	Evidence of the Genre Filter Functional Test	72
Figures 4.23	Evidence of the Recommendation-Card Click Functional Test	72
Figures 4.24	Evidence of the Movie Poster Placeholder Functional Test.....	74

LIST OF TABLES

Table 2.1	Previous Studies on Movie Recommendation Systems ...	8
Table 3.1	Research Dataset Statistics	22
Table 3.2	Proportions and Numbers of Training, Validation, and Test Records	24
Table 3.3	Pseudocode for the Rating- and Genre-Based Recommendation Mechanism.....	29
Table 3.4	Pseudocode for the Recommendation-Card Click Feature..	32
Table 3.5	Example of Users' Average Ratings for the Comedy Genre	33
Table 3.6	Example of Predicted Ratings for Movie A	33
Table 3.7	Example Comparison of Candidate Movies	34
Table 3.8	Project File Structure and Functions	38
Table 3.9	System Development Environment and Libraries	42
Table 3.10	System Functional Testing Scenarios	43
Table 4.1	Dataset Statistics after Preprocessing	46
Table 4.2	Rating Distribution in the Dataset	47
Table 4.3	Distribution of Movies by Genre	49
Table 4.4	Details of the Training, Validation, and Test Data Split....	51
Table 4.5	Summary of Reference User Profiles	52
Table 4.6	Summary of the Model Training Log	54
Table 4.7	Model Evaluation Results on the Test Set	55
Table 4.8	Example of Top-10 Recommendation Results for a Rating of 4.0 and the Comedy Genre	58
Table 4.9	Example of Top-5 Recommendation Results for a Rating of 3.5 and the Drama Genre	60
Table 4.10	Example of Top-5 Recommendation Results for a Rating of 4.0 and the Thriller Genre	62
Table 4.11	Example of Top-10 Recommendation Results for a Rating of 4.5 and the Action Genre	66
Table 4.12	Top-N Selection Testing	68

Table 4.13	Example of Exploration-Mode Recommendation Results Using Duck Soup (1933) as the Reference Movie	69
Table 4.14	Comparison of Recommendations Before and After a Movie Card Is Clicked	69
Table 4.15	Summary of System Functional Testing Results	72

LIST OF SYMBOLS

R	:	User–movie rating matrix of size $m \times n$
r_{ui}	:	Rating assigned by user u to movie i
$\hat{r}_{ui}$	:	Rating predicted by the recommender system for user u and movie i
P	:	User latent-factor matrix of size $m \times k$
Q	:	Movie latent-factor matrix of size $n \times k$
p_u	:	Latent-factor vector for user u
q_i	:	Latent-factor vector for movie i
m	:	Number of users in the dataset
n	:	Number of movies in the dataset
k	:	Number of latent-factor dimensions
λ	:	Regularization parameter used to prevent overfitting
L	:	Loss or error function minimized during model training
K	:	Set of user–movie pairs (u, i) with known ratings
η	:	Learning rate used in Stochastic Gradient Descent (SGD) optimization
$RMSE$	:	Root Mean Square Error, a metric that measures the square root of the average squared difference between actual and predicted ratings
MAE	:	Mean Absolute Error, a metric that measures the average absolute difference between actual and predicted ratings
N	:	Total number of user–movie pairs (u, i) used in the evaluation