**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** ..... **(Advisor I)**
NIP. 19830310 202121 1 006

**Andreas Nugroho Sihananto, S.Kom., M.Kom.** ..... **(Advisor II)**
NIP. 19900412 202406 1 003

**Henni Endah Wahanani, ST. M.Kom.** ..... **(Head Assessor)**
NIP. 19780922 202121 2 005

**Retno Mumpuni, S.Kom., M.Sc** ..... **(Assessor I)**
NIP. 198707162025212045

**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<br>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

## LIST OF FIGURES

## LIST OF TABLES

## 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 |
| $\lambda$ | : | 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 |
| $\eta$ | : | 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 |