# Movie Recommender System

# Tuned Asymmetric Singular Value Decomposition

Winnie Uyen Nguyen | zdnguyen18@earlham.edu

Department of Computer Science | Earlham College, Richmond, IN

## Motivation & Background

As COVID-19 triggered stay-at-home orders and cancelled all social plans, people spend most of their time using online media streaming services. Among those streaming online providers, Netflix became the world's leading Internet television network and the most-valued streaming service, which urges them to personalize users' experience more correctly with the help of recommendation system.

This study designed a movie recommendation system estimating the missing ratings in a user-movie matrix:

- **Model–based Collaborative Filtering:** filtering based on similar taste of users; collecting users' rating on different movies, to create rating matrix for every user and movie.
- **Singular Value Decomposition (SVD):** is suitable for making personalized recommendations based on reactions of other users sharing similar taste

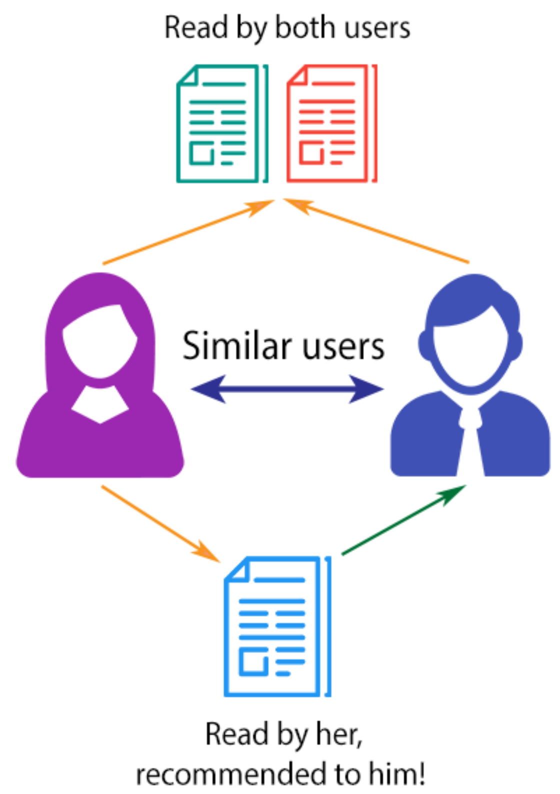



**Figure 1: Collaborative Filtering**

**My major contributions:**

- ✓ Constructing **Asymmetric SVD model** with optimal hyper-parameters to enhance scalability and provide better accuracy
- ✓ Customizing **Estimator** object for the model

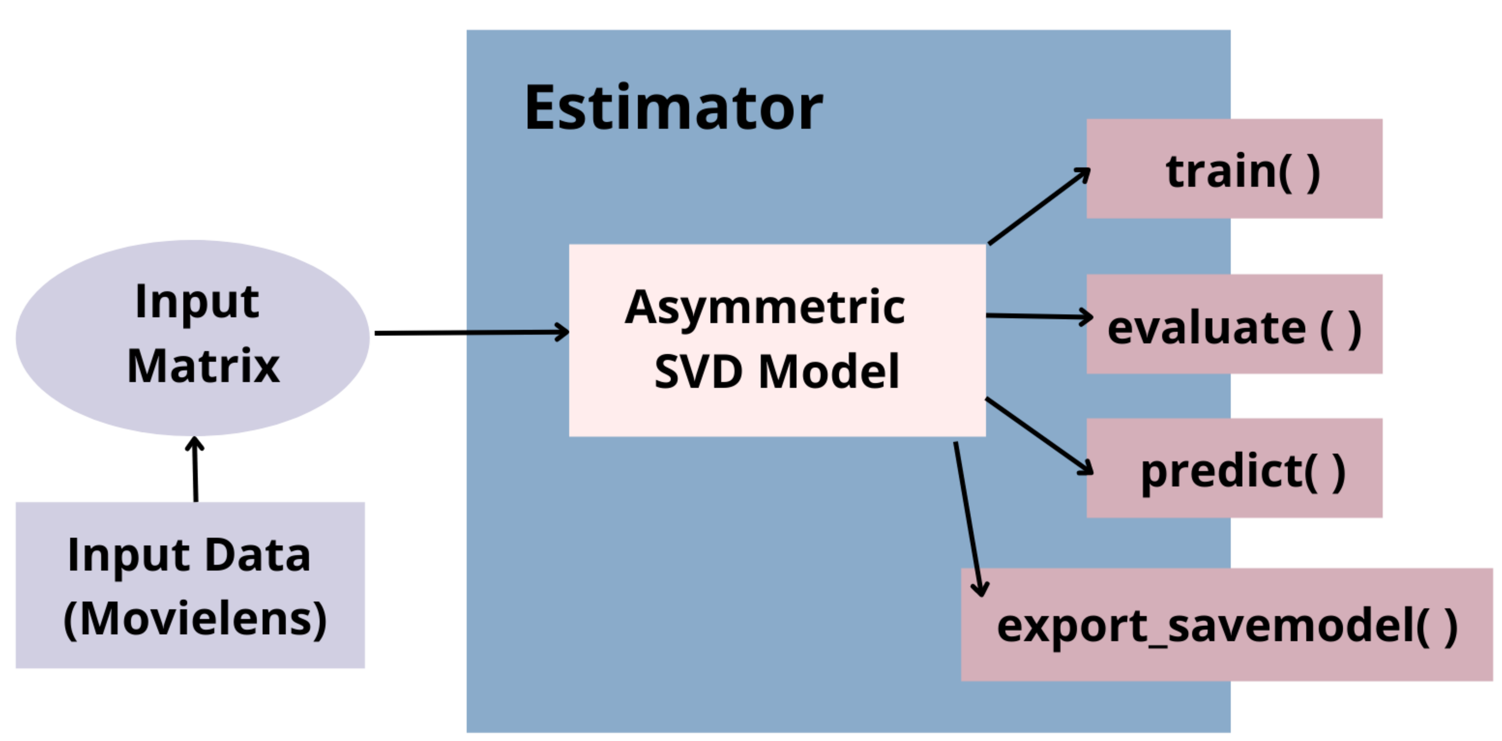



**Figure 2: How the model works**

## Dataset

MovieLens is provided by GroupLens, a web-based research recommender system. 100k Movielens is the oldest dataset, containing a set of movie ratings and demographic information. It has been cleaned so that all users rated at least 20 movies. The rating scale ranges from 1 to 5.

## Methods

**Singular Decomposition Value (SVD):**

- A method based on matrix factorization algorithm to decompose a matrix into 3 small matrices, adding the regularization and bias term.
- **Reduces the size** of original matrix and **optimizes** recommender system.

**Asymmetric SVD with optimal hyper-parameters:**

- ✓ Handle new users with a few ratings without needing to retrain the whole model
- ✓ Having advantages of a model-based algorithm and the ability to take into account both explicit and implicit interactions.

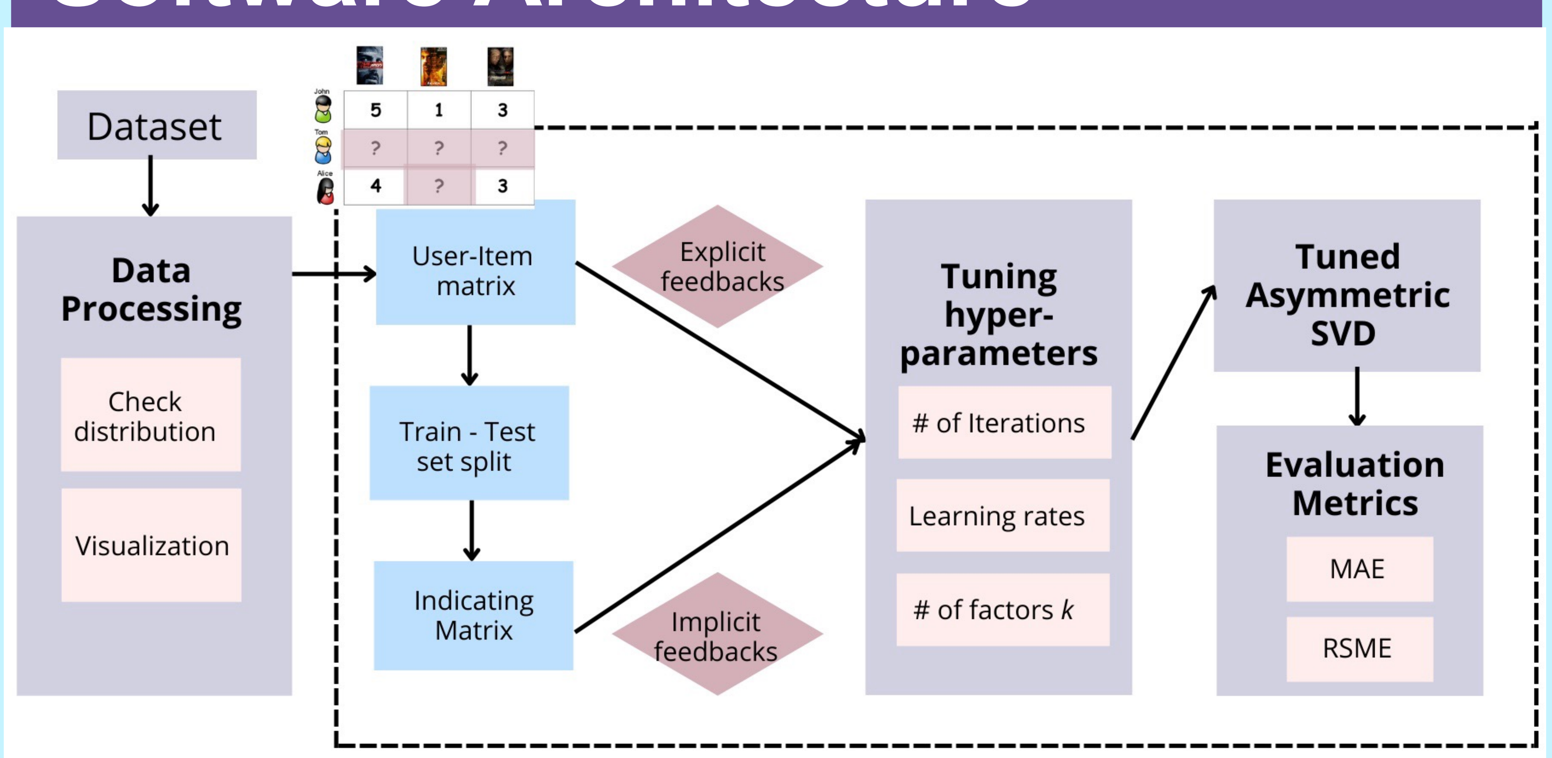
$$\hat{r}_{ui} = b_{ui} + q_i^T\left(p_u + |N(u)| \sum_{j \in N(u)} y_j\right)$$

**Equation 1: Predicted rating**

**Evaluation metrics:**

- Using **Root Mean Squared Error (RMSE)** and **Mean Absolute Error (MAE):** Show how close predicted rating by the algorithm to the true ratings by users

## Software Architecture

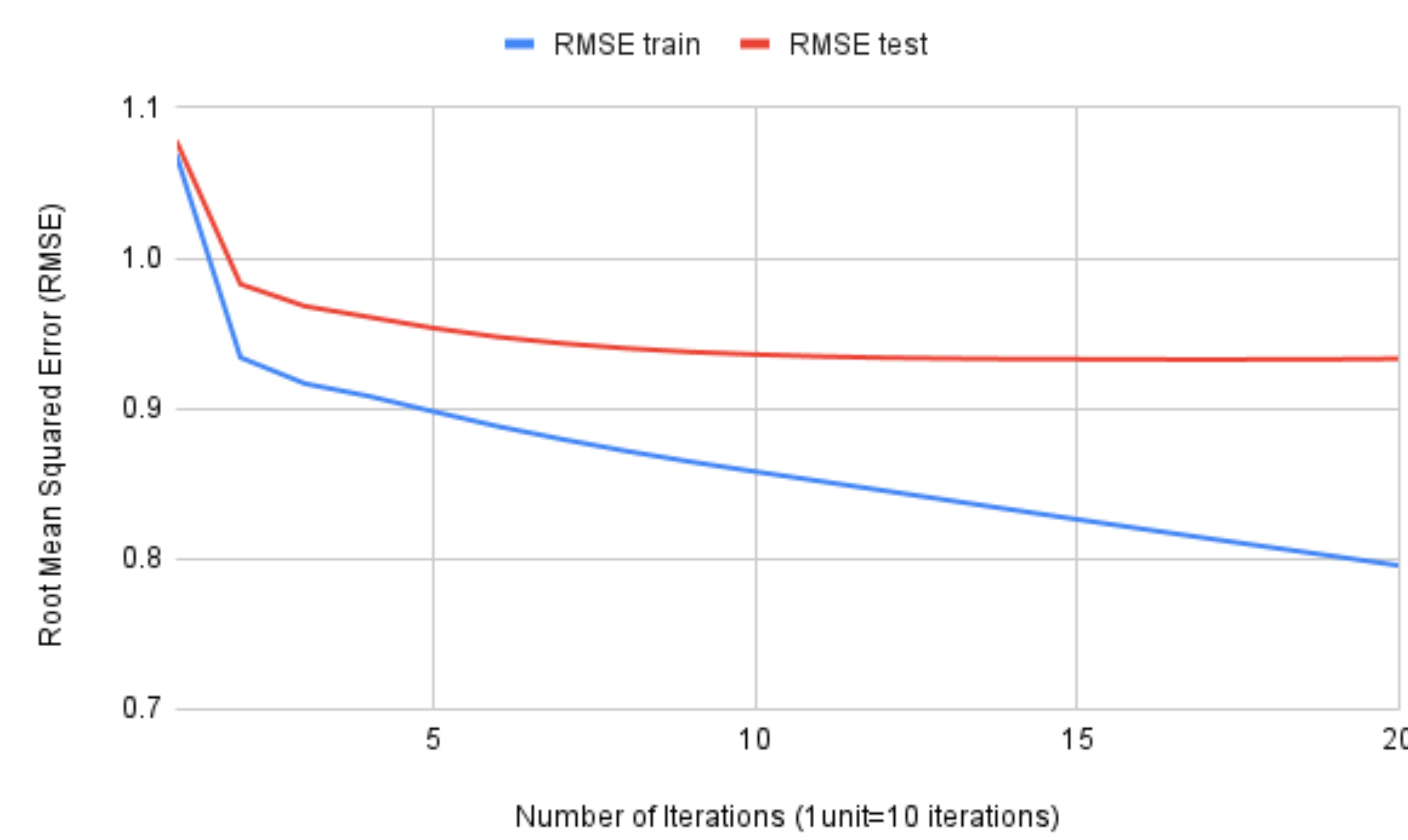



## Result

This study has explored performance of SVD-based algorithms for the movie recommender system. The lower values in both RMSE and MAE for the tuned Asymmetric SVD compared with the regularized version show the better prediction accuracy.

| | RMSE | | MAE | |
|---|---|---|---|---|
| | Train set | Test set | Train set | Test set |
| **SVD** | 0.8912 | 0.9551 | 0.6709 | 0.7535 |
| **Tuned Asymmetric SVD** | 0.6652 | 0.9324 | 0.6257 | 0.7348 |

The chart below captured the RMSEs after the tuned Asymmetric SVD runs 200 iterations. The more iterations we run, the lower value of RMSE we have, meaning the better performance of the recommender system model.




## Future Work

Building Iterative workflow for SVD-based algorithms, working with dynamic dataset: Asymmetric SVD, ApproSVD, Incremental SVD, Appro Incremental SVD

## Acknowledgements

Appreciation to my advisors Dr. Charles Peck and Dr. David Barbella for their helpful feedback and suggestions toward this study.