MOVIE RECOMMENDATION SYSTEM

SAMUEL KARU: GROUP 5
NAMES: Blex Olonde, Dennis Muriungi, Jasho Kiplangat, Shem Omondi and Valary Kones

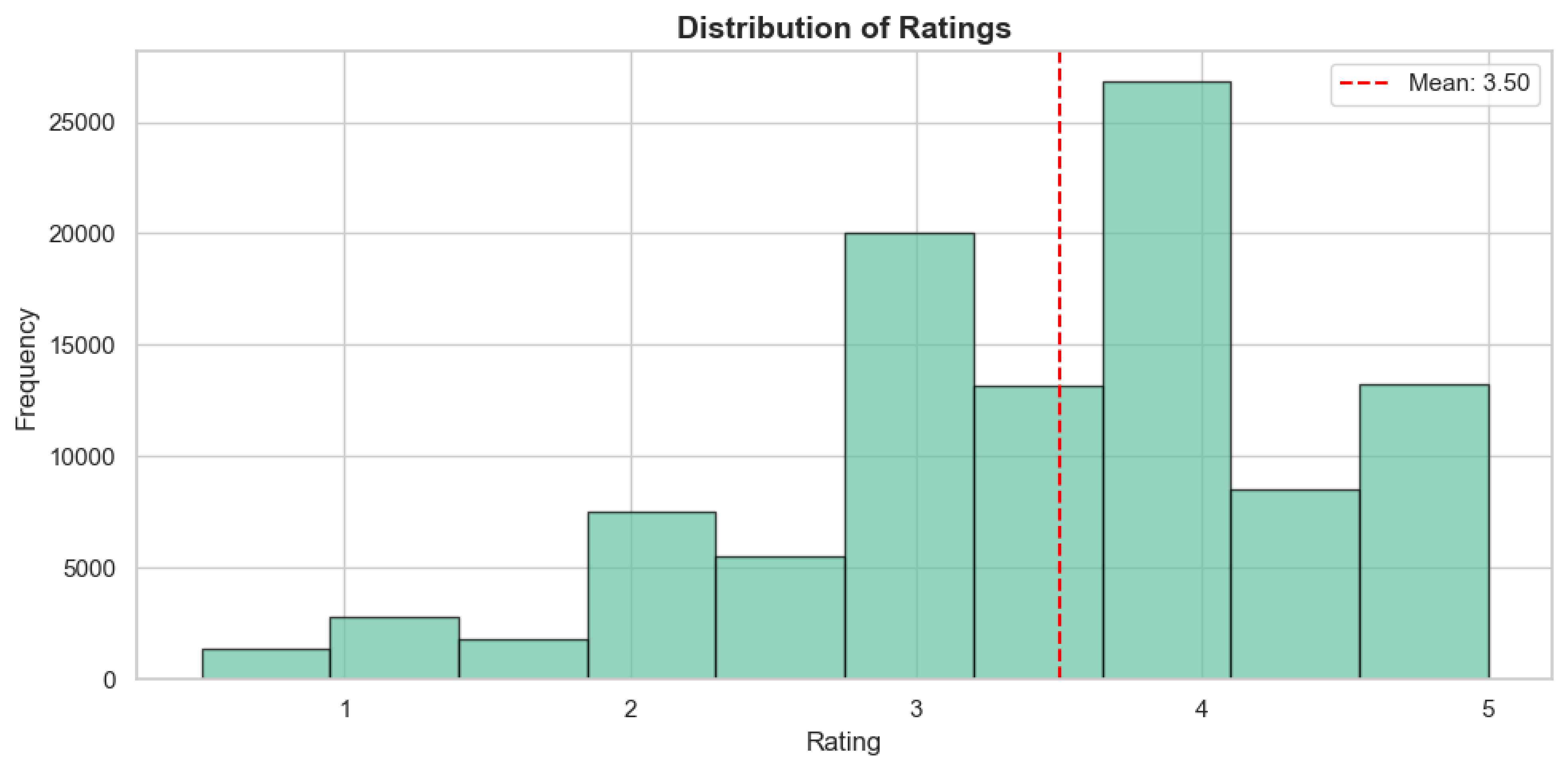
Business Problem

- Users struggle to discover relevant movies
- Large catalogs cause choice overload
- Sparse feedback limits personalization
- Need accurate, scalable recommendation.

Dataset Overview

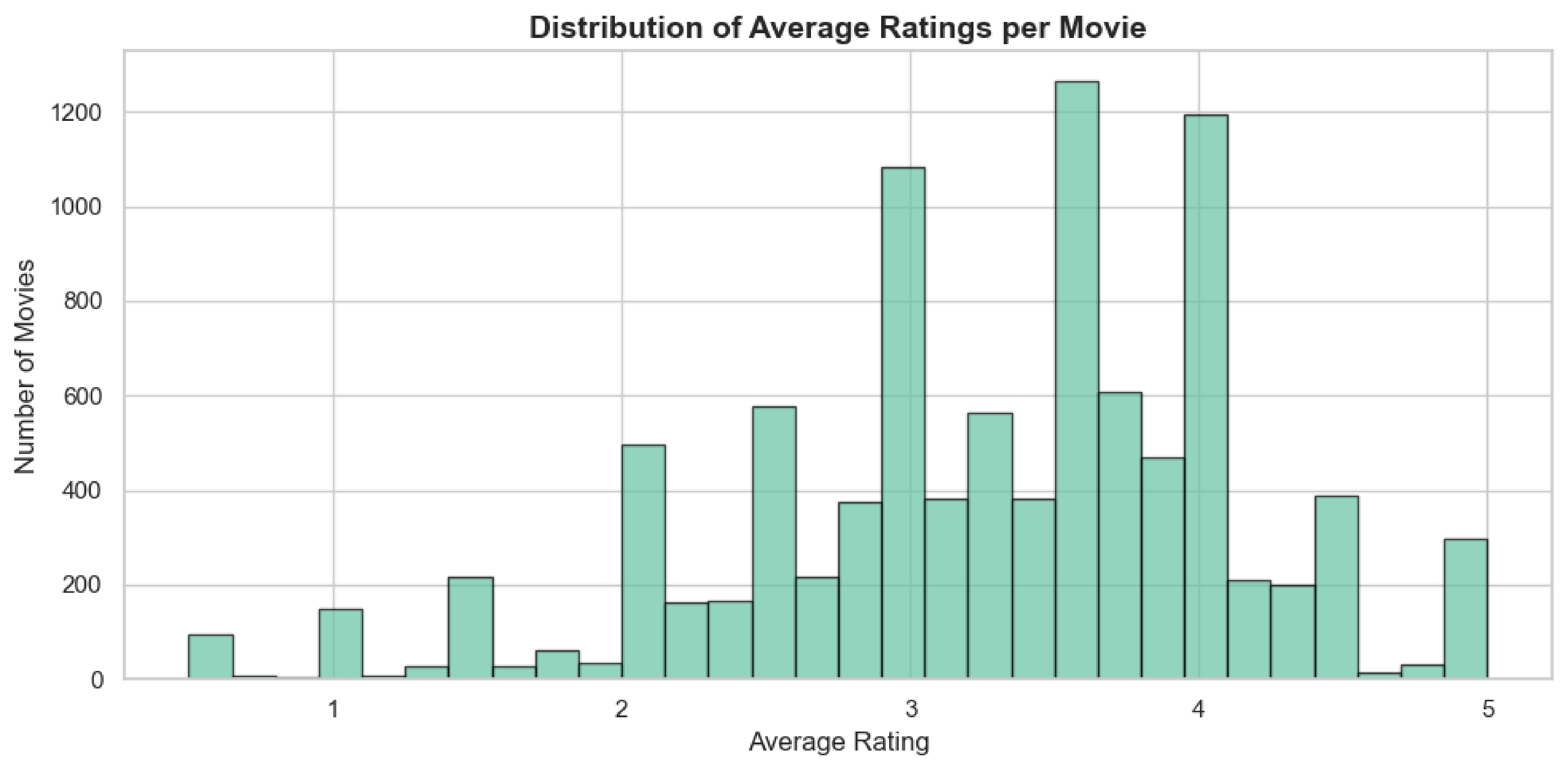
- 100,836 ratings
- 610 users
- 9,742 movies
- 98.3% sparse interaction matrix

Exploratory Insights (Movie Popularity)



- Ratings range from 0.5 to 5.0
- Mean rating = 3.50
- Majority of ratings fall between 3.0–4.0
- Very few ratings below 2.0

Movie Popularity



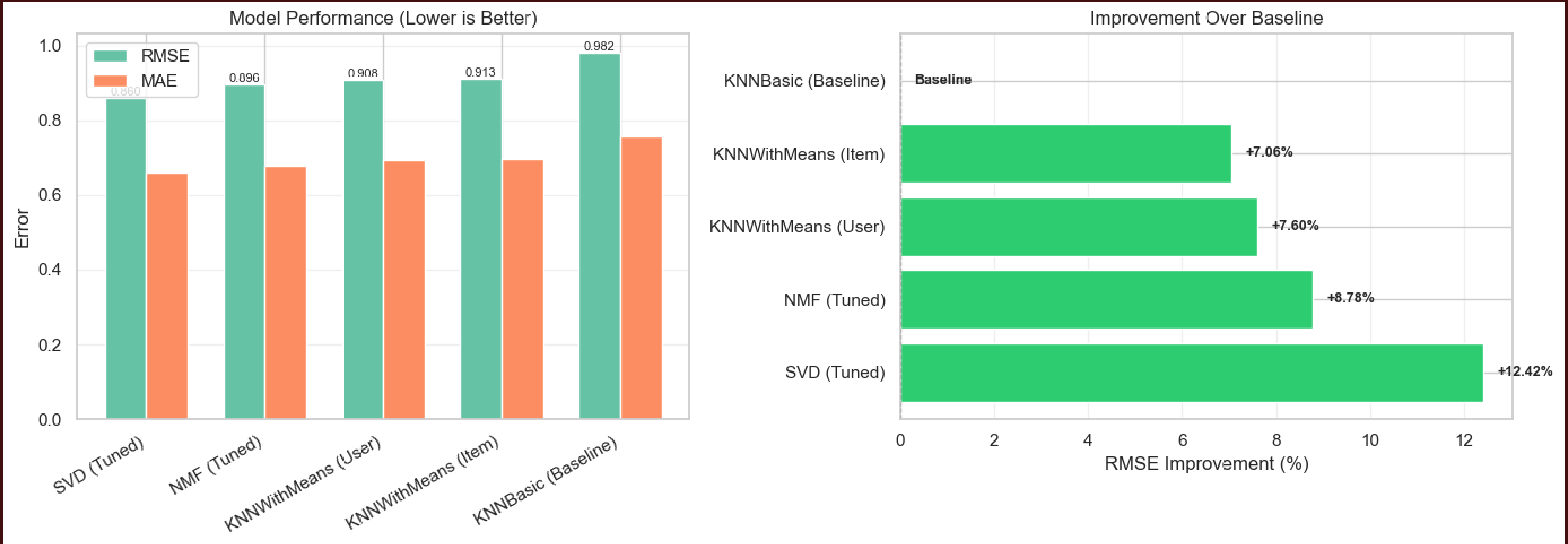
- 7,000+ movies have fewer than 10 ratings
- Only a small number exceed 50 ratings
- Average ratings per movie = 10.4
- Strong long-tail distribution

Modeling Approach

- Baseline: KNNBasic
- Improved KNNWithMeans (user & item)
- Matrix factorization: NMF and SVD
- GridSearch used for tuning

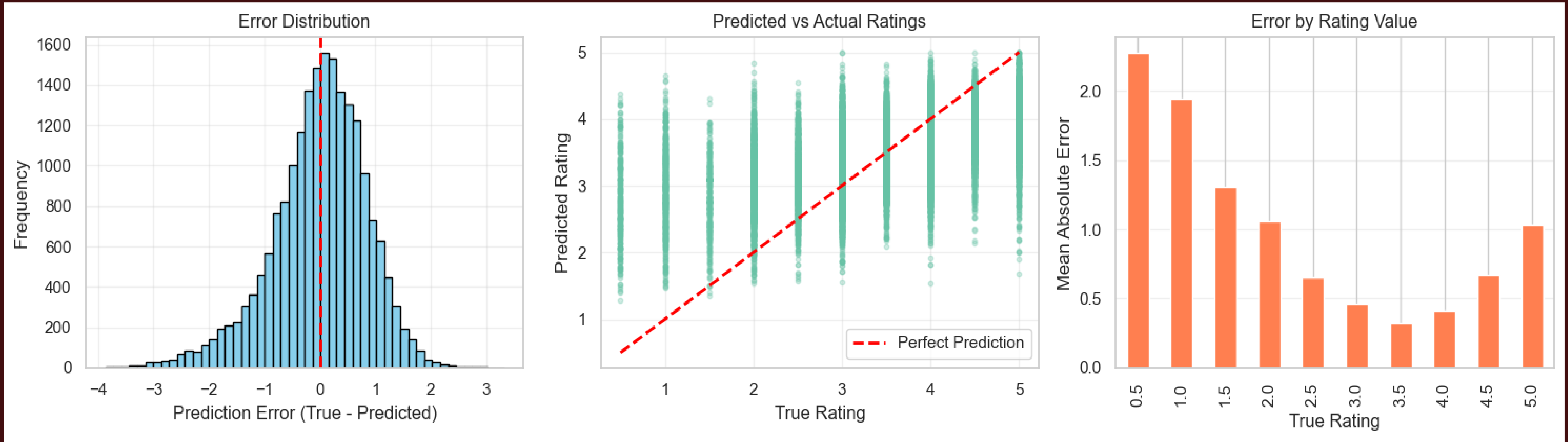
Model Performance

- Baseline RMSE: 0.982
- Best RMSE (SVD): 0.860
- 12.4% improvement over baseline
- SVD selected as final model



Error Analysis (SVD)

- Mean error ≈ −0.016 (unbiased)
- Errors mostly within ±1 rating
- Strong at mid-range ratings
- Weak at extreme ratings



Recommendation Output

- Estimates user preferences for all unseen movies
- Prioritizes titles with the highest predicted ratings
- Delivers a personalized Top-5 movie list
- Avoids recommending previously rated content

THANK YOU