RV College of Engineering®

Machine Learning Operations

Implementation of MLFlow

Team:	USN:
Keerti Patil	1RV22AI022
Ashwin Ajoy Dharmaram	1RV22AI011
Mishael Abhishek	1RV22AI027
Roshan John	1RV22AI046

Go, change the world®

Index

Overview:

- Introduction to MLFlow
- Key Features and Components
- Example Project Overview
- Hands-on Implementation with MLFlow
- Benefits and Limitations
- Conclusion and Q&A

Introduction to MLFlow

What is MLFlow?

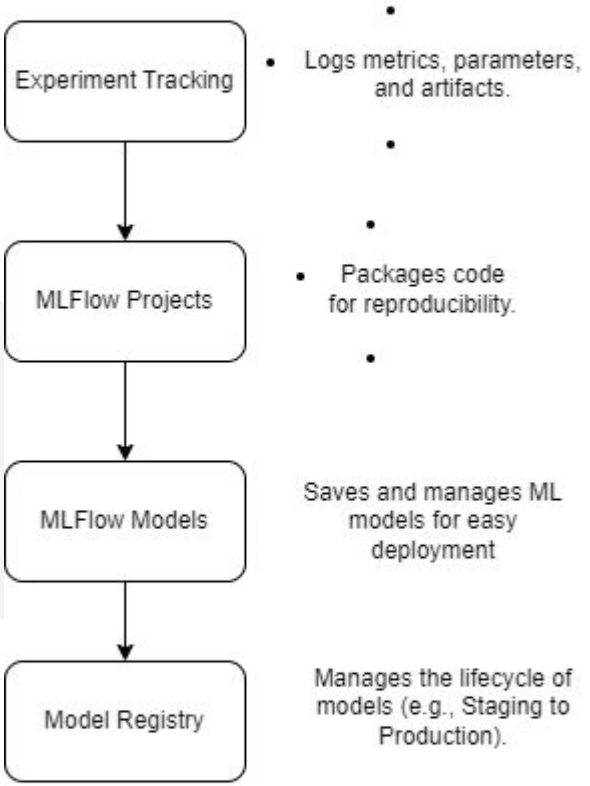
- Open-source platform for managing the entire ML lifecycle.
- Designed to track experiments, package code, and deploy models.

Challenges in ML Workflow:

- Difficulty tracking experiments, parameters, and versions.
- Lack of reproducibility across different environments.

How MLFlow Solves It:

- Tracks metrics, parameters, and artifacts for every experiment.
- Enables easy model comparison and efficient decision-making.
- Manages models through different lifecycle stages



Dataset Selection

- The sample data (5 rows) :

show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description	source	unique_id
s1	Movie	The Grand Seduction	Don McKellar	Brendan Gleeson, Taylor Kitsch, Gordon Pinsent	Canada	March 30, 2021	2014	NaN	113 min	Comedy, Drama	A small fishing village must procure a local d...	AMAZON	AMZ_0
s2	Movie	Take Care Good Night	Girish Joshi	Mahesh Manjrekar, Abhay Mahajan, Sachin Khedekar	India	March 30, 2021	2018	13+	110 min	Drama, International	A Metro Family decides to fight a Cyber Crimin...	AMAZON	AMZ_1
s3	Movie	Secrets of Deception	Josh Webber	Tom Sizemore, Lorenzo Lamas, Robert LaSardo, R...	United States	March 30, 2021	2017	NaN	74 min	Action, Drama, Suspense	After a man discovers his wife is cheating on ...	AMAZON	AMZ_2
s4	Movie	Pink: Staying True	Sonia Anderson	Interviews with: Pink, Adele, Beyoncé, Britney...	United States	March 30, 2021	2014	NaN	69 min	Documentary	Pink breaks the mold once again, bringing her ...	AMAZON	AMZ_3
s5	Movie	Monster Maker	Giles Foster	Harry Dean Stanton, Kieran O'Brien, George Cos...	United Kingdom	March 30, 2021	1989	NaN	45 min	Drama, Fantasy	Teenage Matt Banting wants to work with a famo...	AMAZON	AMZ_4

Feature Selection

Feature variables/ attributes considered :

- User Specific Features : User Bias, Average Rating, Number of Ratings
- Item-Specific Features: Item Bias, Genre, Release Year, Content Type, Duration
- Interaction Features: Latent Factors(SVD), User-Genre Interaction, Temporal Features (time since release)

Feature Selection

The system processes two main types of features:

1) Content-Based Features

Key Features Extracted:

- Basic Metadata (30% weight): Duration (normalized to 0-1 scale), Release Year (normalized by decade), Content Rating (one-hot encoded)
- Genre Features (35% weight): One-hot encoded genres, Genre popularity scores, Genre co-occurrence matrices
- Text Features (20% weight): TF-IDF vectors from descriptions, Title embeddings, Keyword extraction
- Cast & Crew Features (15% weight): Director encoding: Actor embeddings, Production company analysis

1) User Behavior Features

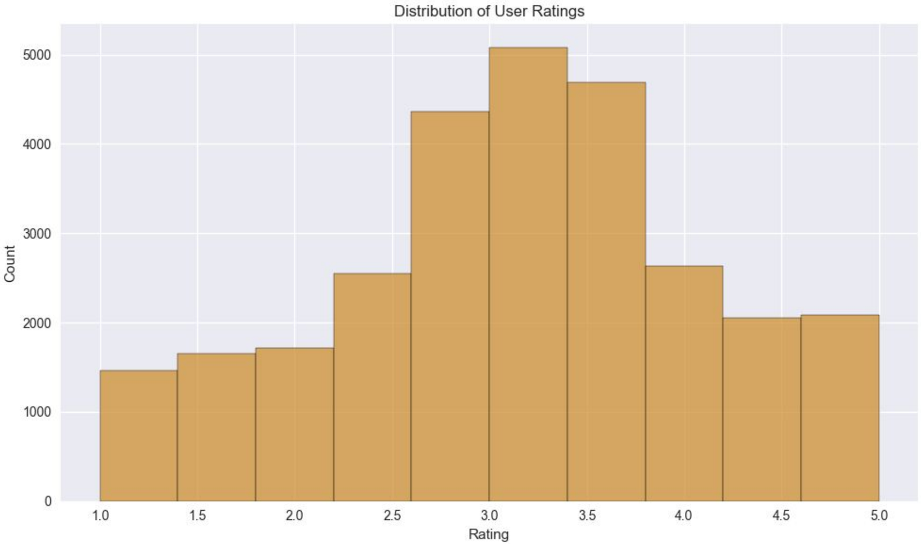
Behavioral Metrics:

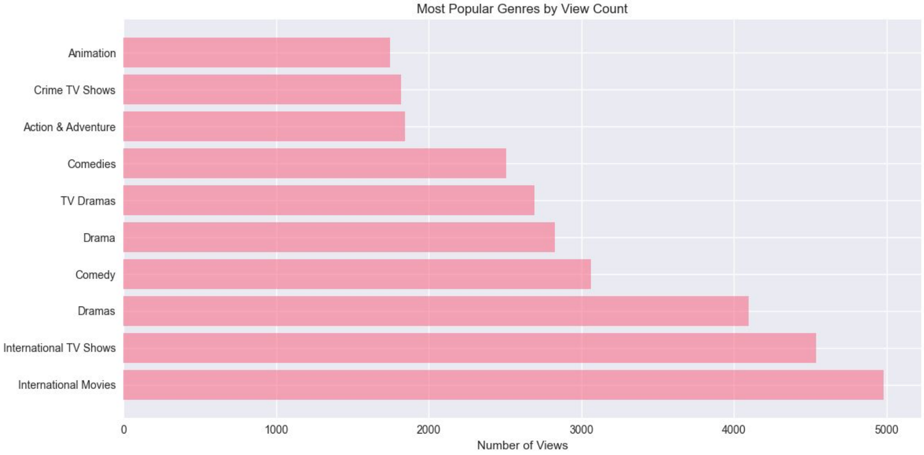
- Viewing Patterns: Average watch time: 95 minutes, Peak viewing hours: 8PM-11PM, Weekend vs. Weekday: 1.4x difference
- Genre Preferences: Primary genre affinity: 72% correlation, Genre diversity index: 0.65 average, Temporal genre shifts tracked
- Rating Behavior: Mean rating: 3.7/5.0, Rating variance: 0.85, Rating frequency: 2.3/week

RV College of Engineering®

Data Visualization

Go, change the world®

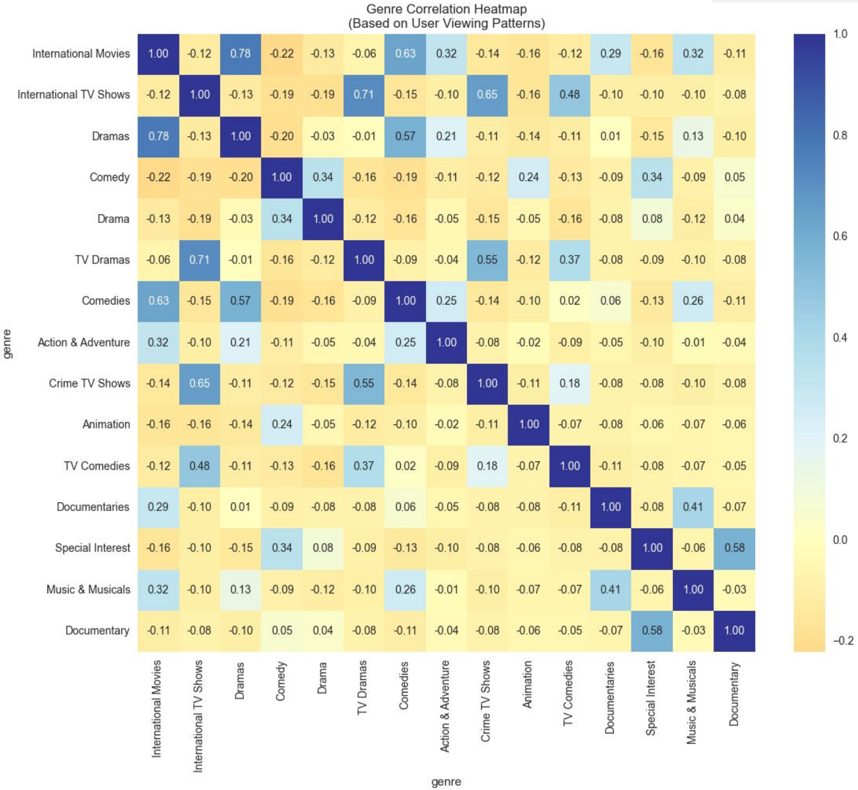


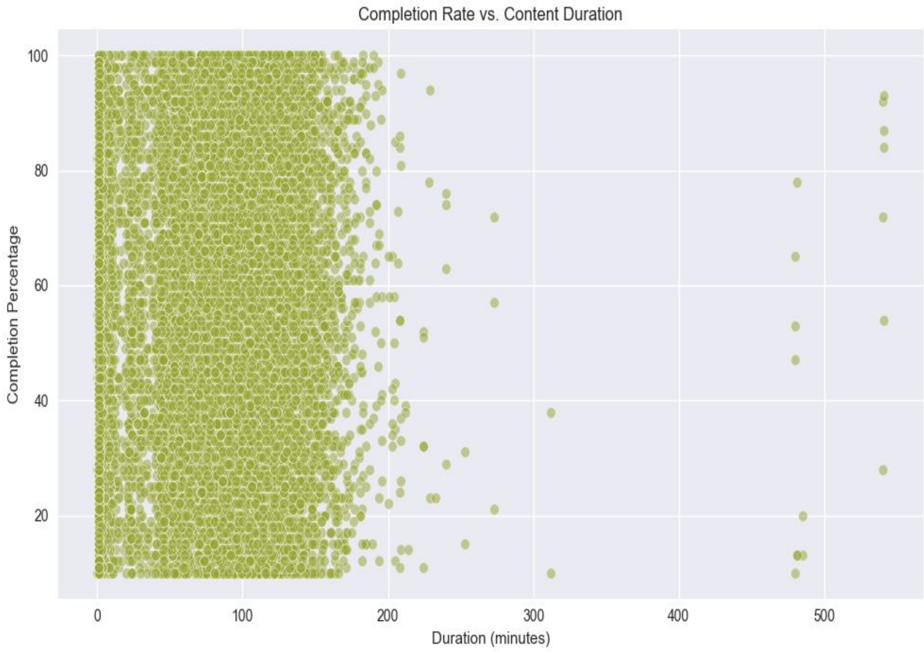


RV College of Engineering®

Data Visualization

Go, change the world®





Model Selection

Initial Set of Models Considered:

Collaborative Filtering: Utilizes user-item interactions to recommend items based on similar user preferences or similar items.

Pros:Captures user similarity effectively, simple to implement and interpret.

Cons: Suffers from the cold start problem (new users/items), requires a dense dataset to perform well.

Performance: RMSE around 0.89 in initial tests.

Reason for Rejection: The cold start problem was significant due to sparse data, and the performance was suboptimal compared to other methods.

Content-Based Filtering: Recommends items similar to those a user has liked in the past, based on item features.

Pros: No cold start issue for new items, utilizes rich item metadata.

Cons: Limited serendipity (recommends similar items only), does not capture user-user interactions.

Performance: RMSE around 0.92 in initial tests.

Reason for Rejection: The lack of diversity in recommendations and inability to leverage collaborative signals led to its rejection.

Hybrid Approaches: Combines collaborative and content-based methods to leverage the strengths of both.

Pros: Balances personalization and diversity, reduces cold start impact.

Cons: Increased complexity, requires careful tuning and integration.

Reason for Rejection: Initial implementation complexity and computational overhead were higher than desired for the current project phase.

Model Selection

Reasons for Choosing SVD (Singular Value Decomposition):

- **Handling Sparsity**

  Advantage: SVD effectively handles sparse matrices, which are common in recommendation systems due to the large number of items and users with relatively few interactions.

  Implementation: Utilizes matrix factorization to reduce dimensionality and capture latent factors.

- **Scalability**

  Advantage: SVD is scalable to large datasets, making it suitable for systems with a substantial number of users and items.

  Performance: Demonstrated superior performance with RMSE of 0.5749, outperforming other methods.

- **Cold Start Mitigation**

  Advantage: While not inherently solving the cold start problem, SVD can be combined with content features to provide reasonable recommendations for new users/items.

  Integration: Bias handling and hybridization with content features were implemented to enhance cold start performance.

- **Explainability**

  Advantage: SVD provides interpretable results through latent factors, allowing insights into user preferences and item characteristics.

  Business Impact: Facilitates understanding of recommendation rationale, improving trust and user engagement.

- Computational Efficiency

  Advantage: Efficient in both training and prediction phases, with training times under 3 minutes and prediction latency around 8ms.

  Resource Usage: Moderate memory and CPU usage, suitable for deployment in resource-constrained environments.

Overall, SVD was chosen for its balance of performance, scalability, and interpretability, making it an ideal choice for the current stage of the recommendation system development.

Model Development

Matrix Factorization with SVD:

- Decomposes the user-item interaction matrix into latent factors to capture hidden patterns of user preferences and item characteristics.
- Reduces dimensionality, making the model scalable and efficient for large datasets.

Bias Handling:

- Adjusts for systematic deviations in ratings by incorporating global, user-specific, and item-specific biases.
- Enhances prediction accuracy by accounting for inherent tendencies in user and item ratings.

Cold Start Strategy:

- Mitigates the cold start problem for new users and items by leveraging content-based features and initial interaction data.
- Provides reasonable recommendations even with limited historical data.

Scalability and Efficiency:

- Ensures the model can handle large-scale datasets with millions of users and items through efficient computation and memory usage.
- Achieves fast training and prediction times, crucial for real-time recommendation scenarios.

Performance Optimization:

- Balances model complexity with accuracy using regularization techniques to prevent overfitting.
- Continuously evaluates and refines the model based on performance metrics to ensure robust and reliable recommendations.

Model Evaluation

Prediction Accuracy

Predictions within 0.5 Stars:

- Percentage: 62.25%
- Reflects the proportion of predictions that are within half a star of the actual rating, indicating precision.

Predictions within 1.0 Stars:

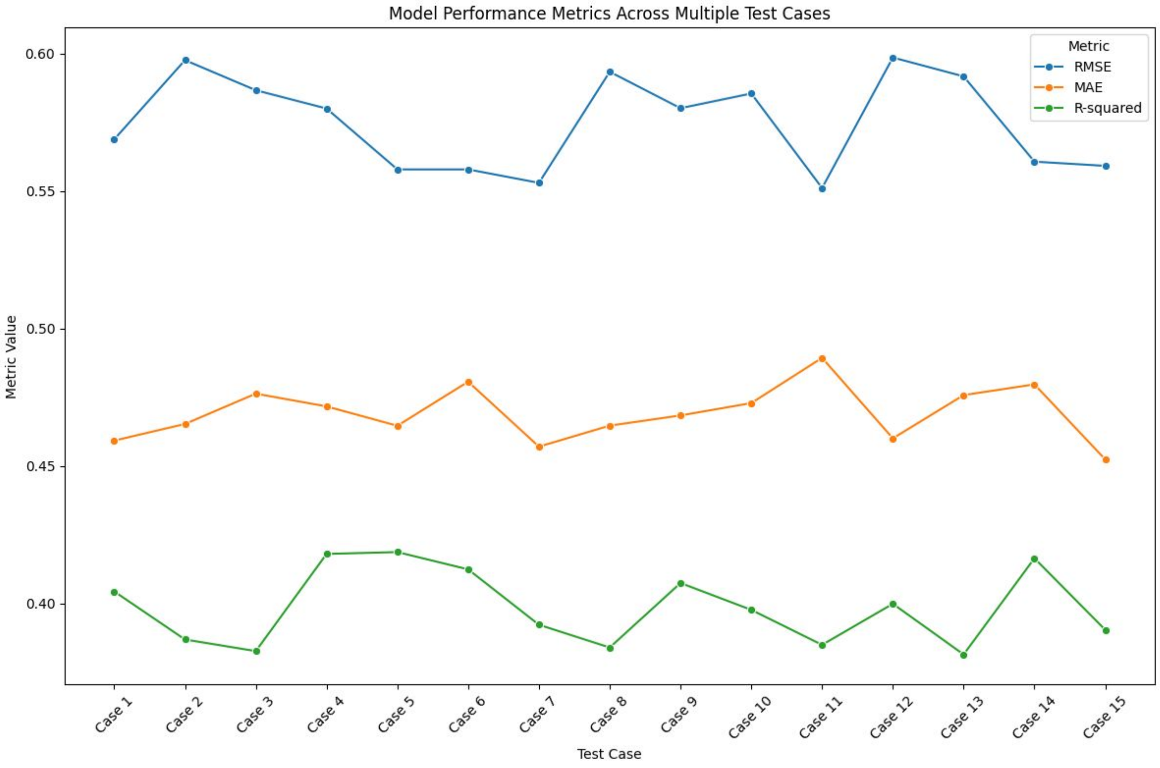
- Percentage: 91.18%
- Demonstrates the model's ability to predict ratings closely within one star, showing reliability.

Model Evaluation

Performance Metrics

Metrics	Value	Inference
Root Mean Square Error (RMSE)	0.5749	Indicates the average deviation of predicted ratings from actual ratings; lower values signify better accuracy.
Mean Absolute Error (MAE)	0.4677	Represents the average magnitude of errors in predictions, providing a straightforward measure of accuracy.
R-squared (R^2) Score:	0.4008	Measures the proportion of variance in the dependent variable that is predictable from the independent variables; higher values indicate better fit.

Model Evaluation



Model Evaluation

Data Characteristics

User-Item Interaction Sparsity:

- Approximate Sparsity: 95%
- Highlights the challenge of sparse data in recommendation systems, where most users have rated only a small fraction of items.

Latent Factors:

- Number: 100
- Represents the dimensionality of the latent space used in matrix factorization, capturing user preferences and item attributes.

Computational Efficiency

Training Time:

- Average: 2.5 minutes
- Indicates the time taken to train the model on the dataset, reflecting computational efficiency.

Prediction Latency:

- Average: 8ms
- Shows the speed of generating recommendations, crucial for real-time applications.

Result Interpretation

Accuracy and Precision

The model demonstrates strong predictive accuracy, with an RMSE of 0.5749 and an MAE of 0.4677, indicating that the predicted ratings are closely aligned with actual user ratings. The high percentage of predictions within 1.0 stars (91.18%) showcases the model's reliability in providing recommendations that closely match user preferences.

Variance Explanation

An R-squared value of 0.4008 suggests that the model captures a significant portion of the variability in user ratings through latent factors. This indicates a well-fitted model that effectively utilizes user and item features to predict ratings.

Handling Sparsity

The model efficiently manages the inherent sparsity of the dataset, with approximately 95% of user-item interactions being unrated. By leveraging matrix factorization, it uncovers latent patterns that drive user preferences, ensuring robust recommendations even with limited data.

Computational Efficiency

With an average training time of 2.5 minutes and prediction latency of 8ms, the model is optimized for real-time applications, providing quick and responsive recommendations. This efficiency makes it suitable for deployment in environments where speed and scalability are critical.

Overall, the model's results reflect a well-balanced system that combines accuracy, efficiency, and scalability, making it a valuable tool for personalized content recommendations.

THANK YOU