Usability Testing Report

Huaye Li

1. Introduction

The seventh week of the Personalized Movie Recommendation System project was dedicated to assessing and improving the usability of the deployed application. The aim of this phase was to conduct systematic testing with real users to evaluate interface intuitiveness, recommendation relevance, and overall user experience. Based on the feedback received, both the AI backend and the user interface were refined to enhance the system's effectiveness, accessibility, and visual clarity. This report outlines the testing methodology, participant feedback, implemented changes, and unresolved issues.

2. Testing Methodology

2.1 Participants

A total of **8 participants** were recruited for usability testing. The sample included:

- 7 graduate students in computer science (experienced users)
- 1 undergraduate students in computer science

The diversity of the group was intentional, reflecting a wide range of digital literacy levels to better assess universal usability.

2.2 Test Environment

- Platform: Streamlit Cloud deployment
- Devices: MacBook Pro (13"), Dell Latitude (15"), and iPad
- Browsers: Chrome, Safari, Firefox
- Access Link: https://ai-movie-recommendation-system-h7ow2ahcckt2tjyyuehxai.streamlit.app

Each session lasted approximately 15–20 minutes and consisted of the following tasks:

1. Enter a valid user ID (between 1–1000)
2. Enter a genre, keyword, or movie title in the search bar
3. Generate and browse recommended movies
4. Interact with movie posters, read overviews, and interpret the relevance

Participants were then asked to complete a structured feedback form and provide open-ended comments.

3. Feedback Summary

Issue Category	Description	Frequency
Text/Input Clarity	Placeholder text not visible enough; unclear what to input	5/8
Button Responsiveness	Button hover state unclear; some users did not realize it was clickable	3/8
Poster Loading Time	Some posters loaded slowly or returned “No Image”	4/8
Contrast Accessibility	Light gray subtitles hard to read on bright screens	6/8
Query Mismatch Handling	No feedback shown when no results are found	3/8
Model Personalization	Users expected the model to reflect past preferences more clearly	2/8

4. Changes Made Based on Feedback

Change Implemented	Description
Improved Placeholder Contrast	Increased contrast ratio of input box placeholders to meet WCAG 2.1 guidelines
Button Styling Enhanced	Added clear hover effects and subtle animation to “Get Recommendations” button
Poster Fallback Logic Updated	Displayed a branded fallback image with text when no poster is found in OMDb
Subtitle Contrast Increased	Changed subtitle font color from #ecf0f1 to a more visible #fdfdfd for better contrast
No Result Feedback	Added alert box using st.warning() to inform users when no matching movies are found
User ID Prompt Clarified	Input label revised to indicate the range and purpose of the ID

These changes were versioned and pushed to the GitHub repository, then redeployed on Streamlit Cloud.

5. AI Model Refinements

While the usability testing focused primarily on front-end interaction, one recurring piece of feedback indicated that the system felt “too generic” in some cases. Upon inspection, it was found that:

- When the search query returned too few movies, fallback logic was not consistently personalized.

Refinement Action:

- The fallback recommendation logic was updated to include the user's top-rated movies from history, even when the search query failed to match content.
- This ensures personalization remains visible, even in cases of limited keyword matches.

6. Remaining Issues and Future Plans

6.1 Remaining Issues

Issue	**Planned Solution**
Query Flexibility	Currently supports keyword search only; future updates will use fuzzy matching (difflib) or NLP-based query parsing
Visual Density on Mobile	Too many cards displayed per row on tablets; needs media query adjustment or st.columns()limitation
Real-Time Feedback Mechanism	No in-app rating or thumbs-up/down feature yet; needed for long-term reinforcement learning and preference learning

6.2 Planned Improvements

- Implement **autocomplete** suggestions for search input using a movie title corpus.
- Add **"like/dislike" interaction** to collect implicit feedback for model retraining.
- Enhance **mobile layout responsiveness** using device-based layout logic or Streamlit's experimental layout tools.

7. Conclusion

The usability testing conducted in Week 7 provided critical insights into the functionality, clarity, and accessibility of the AI Movie Recommendation System. As a result, both visual and functional aspects of the interface were improved, and minor refinements to the recommendation logic were made. While some challenges remain—particularly with respect to query flexibility and mobile responsiveness—the overall system now offers a more robust, user-friendly experience aligned with real-world user expectations.