Conceptualized User Study Documentation✍️

---

1. Hypothesis 💡

The central hypothesis for this study is that **users who receive contextualized impact information about how their movie ratings affect the recommendation system will report a significantly higher level of perceived accuracy and trust in the recommendation system** compared to users who receive no such information.

This hypothesis is based on the fact that,when a system is transparent about how it learns, users feel more in control. This leads them to give better input, and in turn, they see the system as more effective and trustworthy.

---

2. Study Design: A/B Testing (Between-Subjects)⚖️

To test this hypothesis, an A/B testing, or **between-subjects design**, is implemented. This involves two distinct groups of participants who will interact with different versions of the system.

- **Group A (Experimental Group)**: This group interacts with the full, guided active learning system. After each movie rating, they will be shown an **impact message** that explains how their rating (e.g., on a new genre or a familiar one) helps the system learn about their tastes.
- **Group B (Control Group)**: This group interacts with a version of the system where the **impact message** is absent. They will follow the same rating process but will receive no real-time feedback on how their input affects the system.

By comparing the data collected from both groups, the isolated effect of the **impact message** can be measured.

---

3. Participants and Recruitment Process 👥

- **Target Audience**: The study targets a diverse group of general adult internet users. The participants should be "new users" to the system, meaning they have no prior history with the recommender.
- **Sample Size**: A minimum of 30-50 participants per group is required to achieve a statistically significant result, though a larger sample size would increase the statistical power of the findings.
- **Recruitment Channels**: Participants could be recruited from academic platforms, social media, or professional networks.
- **Incentives**: To encourage participation and ensure a good response rate, a prize draw entry is offered as compensation.

---

4. Study Procedure📝

The study is designed to be a seamless, multi-step process for the user:

1. **Welcome and Informed Consent**: Upon starting the study, the user is presented with a dedicated **consent page**. This page explains the study's purpose, the tasks involved, and ethical considerations regarding data privacy and voluntary participation. The user must explicitly agree to the terms to proceed.
2. **Random Assignment**: After consenting, the system automatically and randomly assigns the participant to either **Group A** or **Group B**. This assignment is stored in the session data and is hidden from the user to prevent bias.
3. **Cold-Start Rating Phase**: The user is guided to the interactive rating interface. They are asked to rate 10 randomly selected movies, one at a time. The interface provides instructions to rate from 0.5 to 5.0, with an option to rate their interest level for movies they have not watched.
   - **Group A**: Receives the dynamic **impact message**.
   - **Group B**: Completes the same process but without the **impact message**.
4. **Recommendations Display**: After rating all 10 movies, the system generates and displays a list of personalized movie recommendations based on the user's ratings.
5. **Post-Study Questionnaire**: Following the recommendations page, the user is automatically redirected to a **final questionnaire**. This questionnaire is crucial for testing the hypothesis. It includes:
   - **Quantitative Metrics**: Likert scale questions (1-5) to measure the user's **perceived accuracy** of the recommendations and their **trust** in the system.
   - **Qualitative Feedback**: An open-ended text box for users to provide comments on their experience.
6. **Debriefing and Compensation**: The user is directed to a final **debriefing page**. This page thanks them for their participation, explains the true purpose of the study (comparing the two groups), and provides details on how they will be compensated.

---

5. Data Collection and Analysis📊

- **Data Points**: The system collects the user's group assignment, their 10 ratings, and their responses from the final questionnaire. This data is stored in a dedicated database model.
- **Statistical Analysis**: The primary analysis would involve running **independent samples t-tests** to compare the mean perceived accuracy and trust scores between Group A and Group B. A statistically significant difference would support the hypothesis.
- **Qualitative Analysis**: The open-ended feedback would be reviewed for recurring themes and insights, particularly from Group A, to understand their experience with the impact messages.
