Hybrid Book Recommendation System

1. Overview of Collaborative and Content-Based Filtering

Collaborative Filtering:

- **Strengths:**
  - Learns from actual user preferences.
  - Effective for discovering hidden patterns and unexpected recommendations.
- **Weaknesses:**
  - Suffers from the cold-start problem (new users or items).
  - Requires sufficient user-item interaction data.

Content-Based Filtering:

- **Strengths:**
  - Works well for new users or items.
  - Uses item metadata like author, genre, etc., for personalized suggestions.
- **Weaknesses:**
  - May lead to over-specialization (lack of novelty).
  - Quality depends on the completeness of metadata.

2. Chosen Hybrid Strategy: Blending (Weighted Average)

- To leverage the strengths of both methods, we used the **blending strategy**.
- The final recommendation score is computed using a weighted average:

  $$\text{Final Score} = \alpha \times \text{Collaborative Score} + (1 - \alpha) \times \text{Content-Based Score}$$

- The blending factor α (alpha) is tunable (e.g., 0.5 for equal weight, 0.7 for more emphasis on collaborative filtering).

3. Logic and Implementation

- Input: User ID and a Book Title.
- Output: Top-N recommended books with scores and source attribution.
- If the book title is provided, content-based similarity is computed using TF-IDF on combined metadata (title, author, publisher).
- If the user has sufficient history, collaborative filtering suggestions are fetched.

- Scores are normalized and blended to produce final recommendations.

4. Sample Outputs (Screenshots)

- 1: Recommendations for an active user

```
Top-N Hybrid Recommendations:
Book: 0743454529, Score: 4.551127890215301, Source: Hybrid
Book: 0345339738, Score: 4.5359279727324004, Source: Hybrid
Book: 0618002235, Score: 4.513282479802655, Source: Hybrid
Book: 0836220889, Score: 4.508350642169669, Source: Hybrid
Book: 0439425220, Score: 4.500839790400001, Source: Hybrid
Book: 0439136369, Score: 4.47456156797508, Source: Hybrid
Book: 0836213319, Score: 4.472428006462371, Source: Hybrid
Book: 0345348036, Score: 4.471106184911133, Source: Hybrid
Book: 067168390X, Score: 4.454391504152559, Source: Hybrid
Book: 0553274325, Score: 4.448115037594795, Source: Hybrid
```

- 2: Recommendations for a new user with only content-based suggestions

```
Content-Based Recommendations for New User 999999:
1. 0439425220 | Score: 4.5288 | Source: Hybrid
2. 0743454529 | Score: 4.5226 | Source: Hybrid
3. 067168390X | Score: 4.5083 | Source: Hybrid
4. 0330262130 | Score: 4.4832 | Source: Hybrid
5. 0345339738 | Score: 4.4781 | Source: Hybrid
```

5. Interpretation of Results

- Active Users: Benefit from collaborative filtering, especially when they have rated many books.
- New Users: Rely more on content-based recommendations until enough data is available.
- Hybrid System: Balances diversity and personalization effectively by combining both sources.

6. Observations and Suggestions for Tuning

- **α Value**:
  - Higher α improves recommendations for active users.
  - Lower α helps cold-start users.
- **Possible Improvements**:
  - Context-aware hybridization (e.g., using time, mood, or location).
  - Temporal dynamics: Account for changing user preferences over time.
  - Incorporating implicit feedback like clicks, views, or time spent on a book page.