1. Summary of Tuned Parameters

Collaborative Filtering (User-User):

- K (number of neighbors): [5, 10, 20, 30]

- Similarity Threshold: [0.0, 0.1, 0.2]

- Metric Evaluated: Precision@5

Content-Based Filtering (TF-IDF):

- ngram_range: [(1, 1), (1, 2)]

- max_features: [1000, 5000, 10000]

- min_df, max_df: (currently not tuned, future work)

- Metric Evaluated: Precision@5

2. What Worked and What Didn't

Collaborative Filtering (CF):

- Best performance was observed for:

  - K=30, Threshold=0.2 → **Precision@5 = 0.771**

- Higher K generally improved performance.

- Adding a threshold (> 0.2) helped filter weak similarities.

Content-Based Filtering (TF-IDF):

- Precision remained **low overall**.

- Best TF-IDF config:

  - ngram_range=(1, 2), max_features=5000 → **Precision@5 = 0.048**

- Bigger max_features didn't consistently help.

- ngram_range=(1, 2) was slightly better than (1, 1).

3. Visual Comparison

Collaborative Filtering Heatmap:

- Best score: Top-right corner

- Precision@5 increased with larger K and higher thresholds

Content-Based Filtering Bar Plot:

- Grouped by (ngram_range, max_features)

- Tiny precision gains across configurations

- Lower consistency compared to CF

4. Final Selection Rationale

Method	Final Parameters	Precision@5
User-User CF	K=30, Threshold=0.2	**0.771**
Content-Based	ngram=(1, 2), max_features=5000	0.048

- **User-User CF** is selected as the **primary recommender**.
- **Content-Based** can assist with **cold-start users** or new items.

5. Manual Observations

- **User-User CF** handled known users well.
- **Cold-start issue** persists in CF when users have <5 ratings.
- Content-based can still generate basic recommendations in those cases, but quality is low.