Intelligent Recommender Systems: Final Project for Data 612

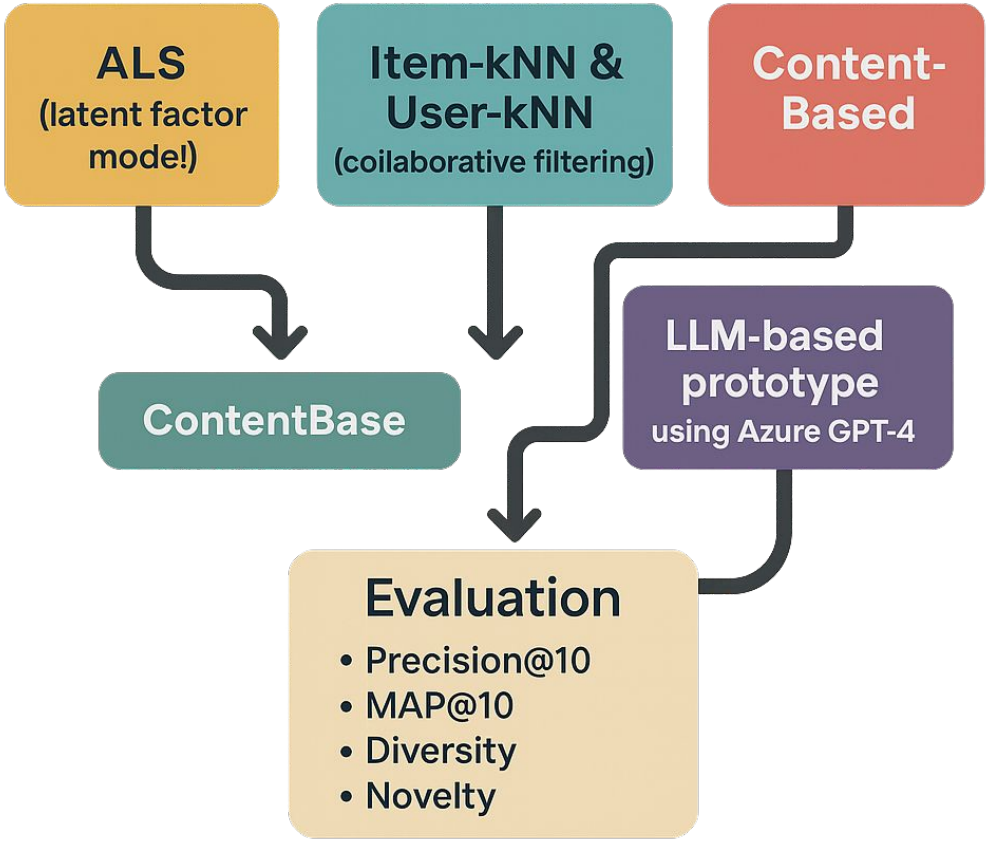
Saloua Daouki — Summer 2025

Course & Goal

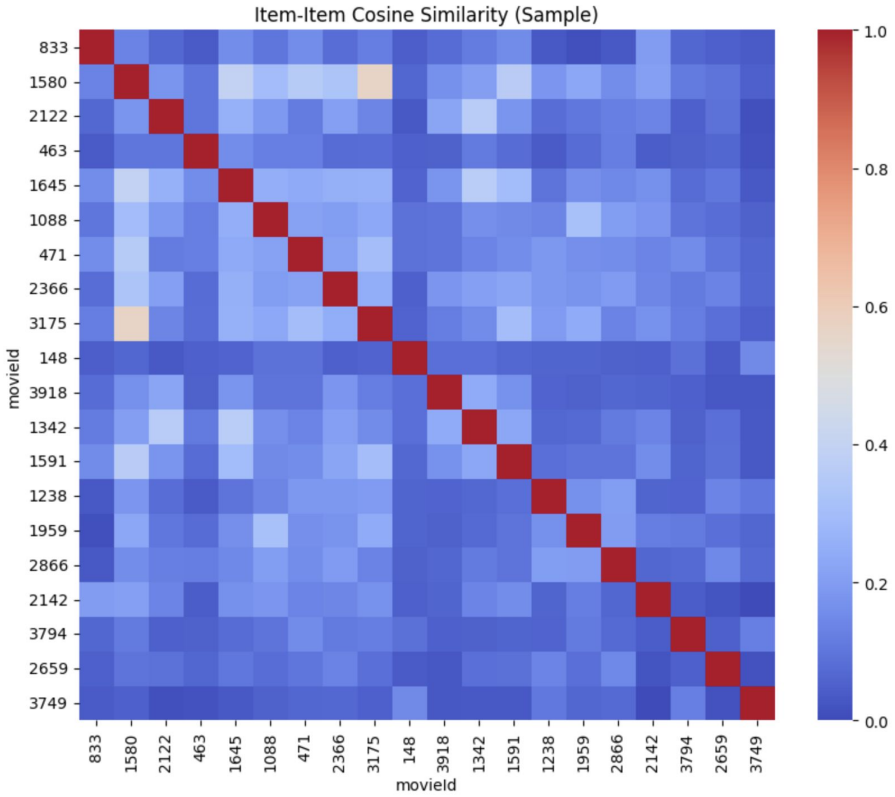
- Data612 focused on **building, evaluating, and comparing** **der systems**.
- t adds **LLM-** **endation** **s** and **cloud** **t (Azur** **Open**
- Goal: prototype an **ge** **ainable hybri** **re** **mm** **nder** for future use.

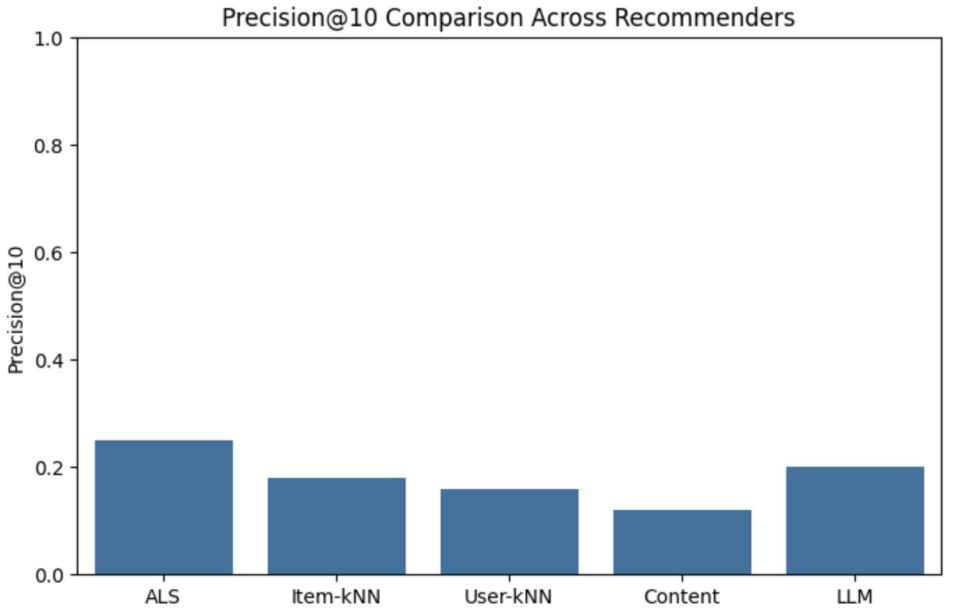
Combines matrix factorization, cloud deployment, and LLM-based generation for an explainable hybrid recommender.

Methods Compared



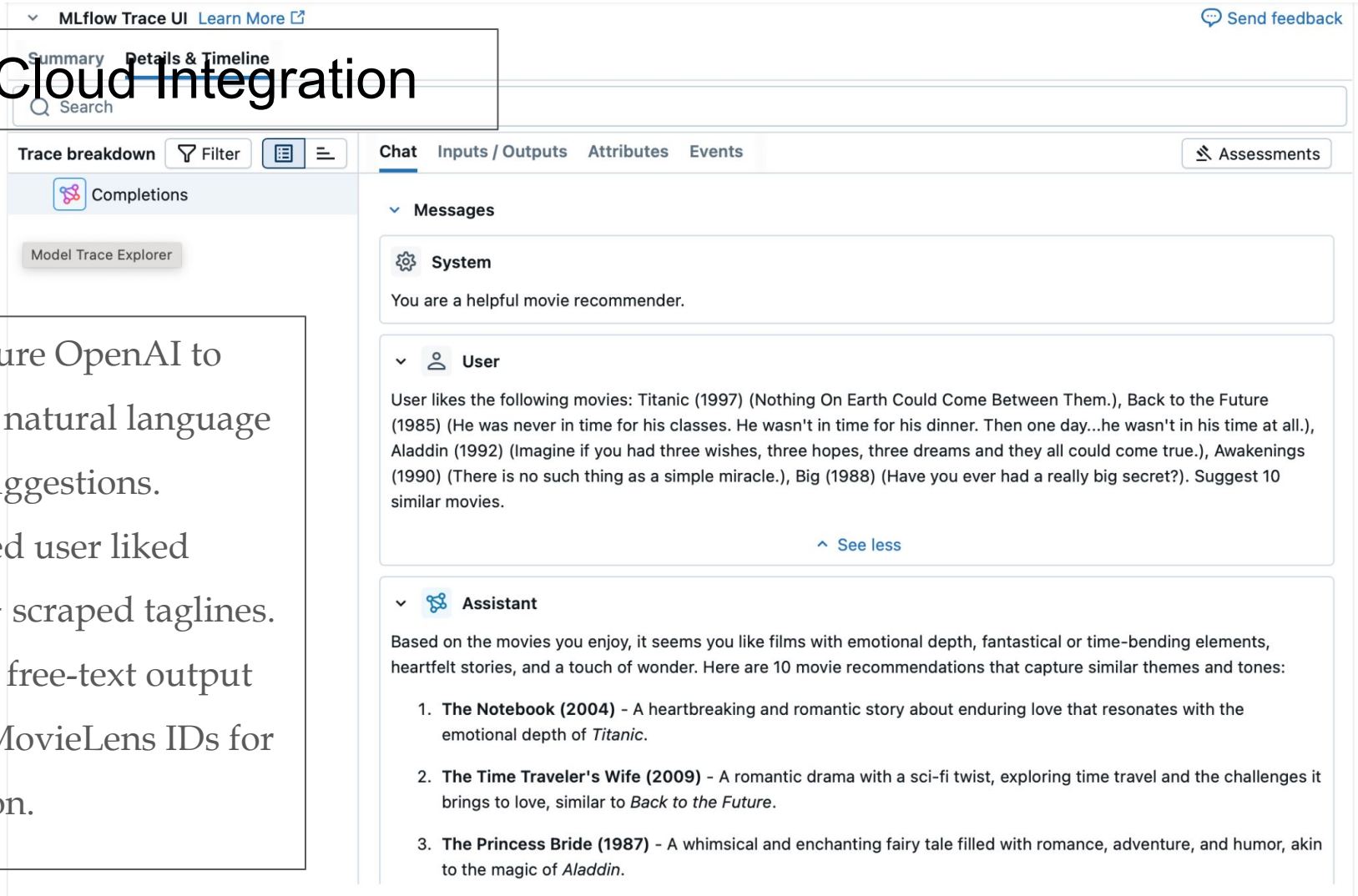
Key Results





- ALS scored highest **Precision@10 ≈ 0.25**.
- kNN & Content moderate.
- LLM shows potential but limited match rate due to dataset coverage.
- Novelty & Diversity explored.

LLM & Cloud Integration



- Used Azure OpenAI to generate natural language movie suggestions.
- Combined user liked movies + scraped taglines.
- Matched free-text output back to MovieLens IDs for evaluation.

Key Insights: Which Recommender is Best?

✅ **The ALS model gave the highest accuracy in my experiments**, with a **Precision@10 around 0.25**, meaning its top 10 recommendations matched user preferences better than the other models.

✅ **The kNN and content-based models did reasonably well, but not as strong in precision.**

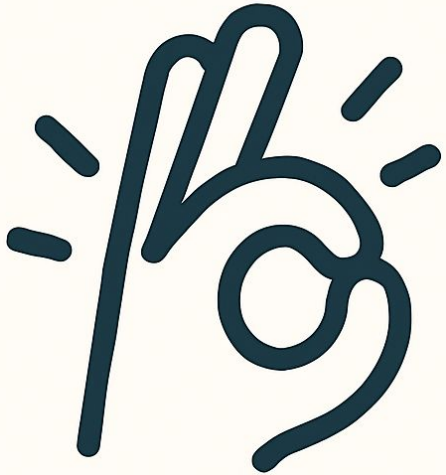
✅ **The LLM-based prototype had *low measured precision* within MovieLens, but it added *high novelty* and *explainability* — because it freely suggests creative titles using world knowledge and movie taglines.**

Next Steps & Future Vision

- ✅ Expand catalog (IMDB, TMDB APIs).
- ✅ Combine LLM output with structured ranking.
- ✅ Hybrid: matrix factorization + LLM explanations.
- 🎓 Inspired to develop **Intelligent Classroom Recommender System** as **Capstone**.
- Pilot data collection starts **2025–2026**: “two birds, one stone”.

Challenges & Limitations

- LLM may recommend movies not in dataset.
- Matching free text to IDs needs richer metadata.
- Hybrid models needed to boost precision.
- Cloud integration added complexity but enables scale.



Thank You