Evaluating Graph Transformers for Scalable Social Media Recommendations

Aleck Wu a5wu@ucsd.edu
SangGyu An sgan@ucsd.edu
Mentor: Yusu Wang yusuwang@ucsd.edu
Mentor: Gal Mishne gmishne@ucsd.edu

UC San Diego
HALICIOĞLU DATA SCIENCE INSTITUTE

Background

What is Bluesky?
A decentralized social media platform where user data is distributed across multiple servers rather than a single central database. This design aims to give users more control but makes traditional recommendation approaches more challenging.

- **No centralized engagement data** – User interactions are distributed across independent servers.
- **Lack of personalization** – Popularity-based methods ignore individual user interests.
- **Scalability** – Large graphs make real-time inference costly.

Why is this important: Decentralized platforms are growing, but personalized content discovery remains an open challenge. An effective recommender system can enhance user engagement and improve the decentralized social media experience.

Objective

We aim to develop a scalable recommendation pipeline for Bluesky that efficiently captures user interests in real-time while balancing **personalization** (measured by **MRR**) and **content diversity** (measured by **ILD**).

Data

Dataset Overview
This study examines interactions from 2023-6-7 to 2023-6-14, capturing user activity such as posts, likes, and follows.

Statistic	Count
Users	23,765
Posts	365,314
Likes	1,042,739
Follows*	7,301,917

Table 1: * Follows data from 2023-1-1 to 2023-6-30

Visualizing Bluesky Data
Below, we illustrate the Bluesky homepage (left) and a graph representation of user-user and user-post interactions (right), where nodes represent different users and posts.

Fig. 1: Bluesky homepage

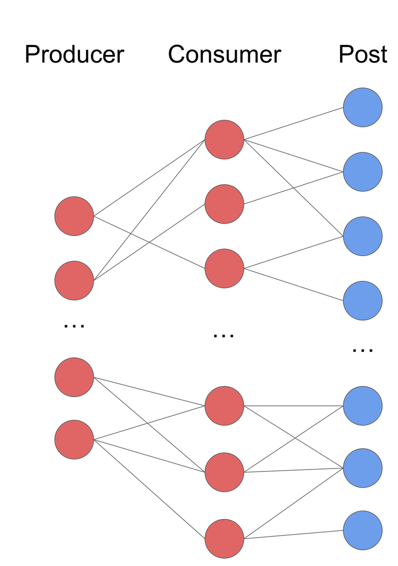

Fig. 2: Interaction visualization

Methodology

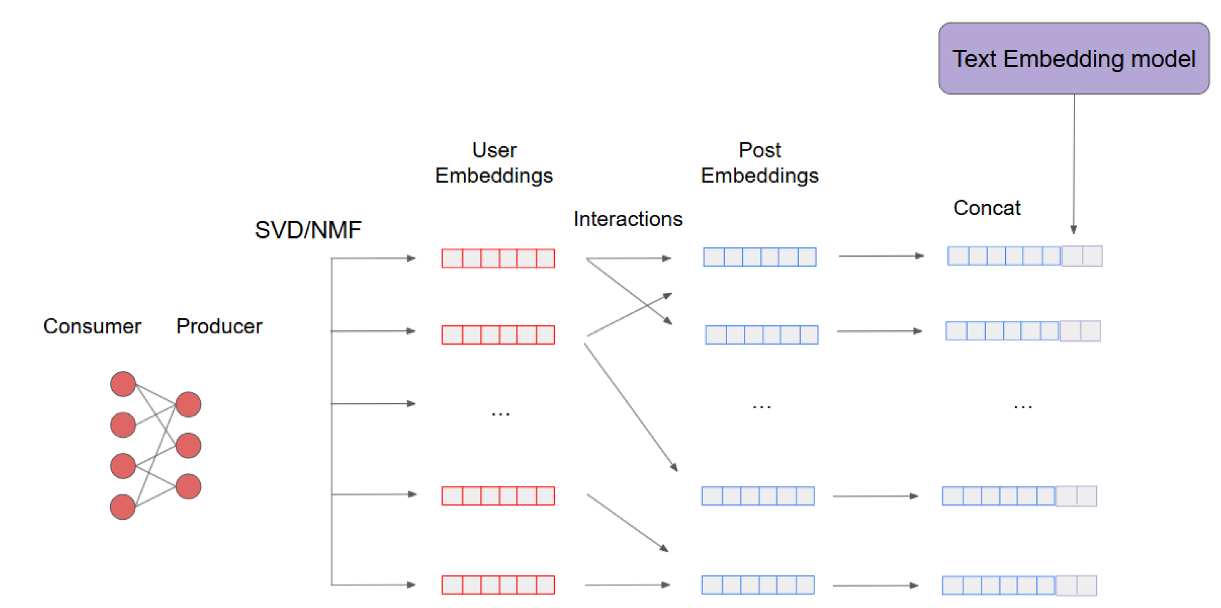

Fig. 3: User and Post Embeddings Generation Process

- **User Embeddings:** constructed from the consumer-producer bipartite follow graph.
- **Post Embeddings:** constructed in real-time by aggregating user interactions.

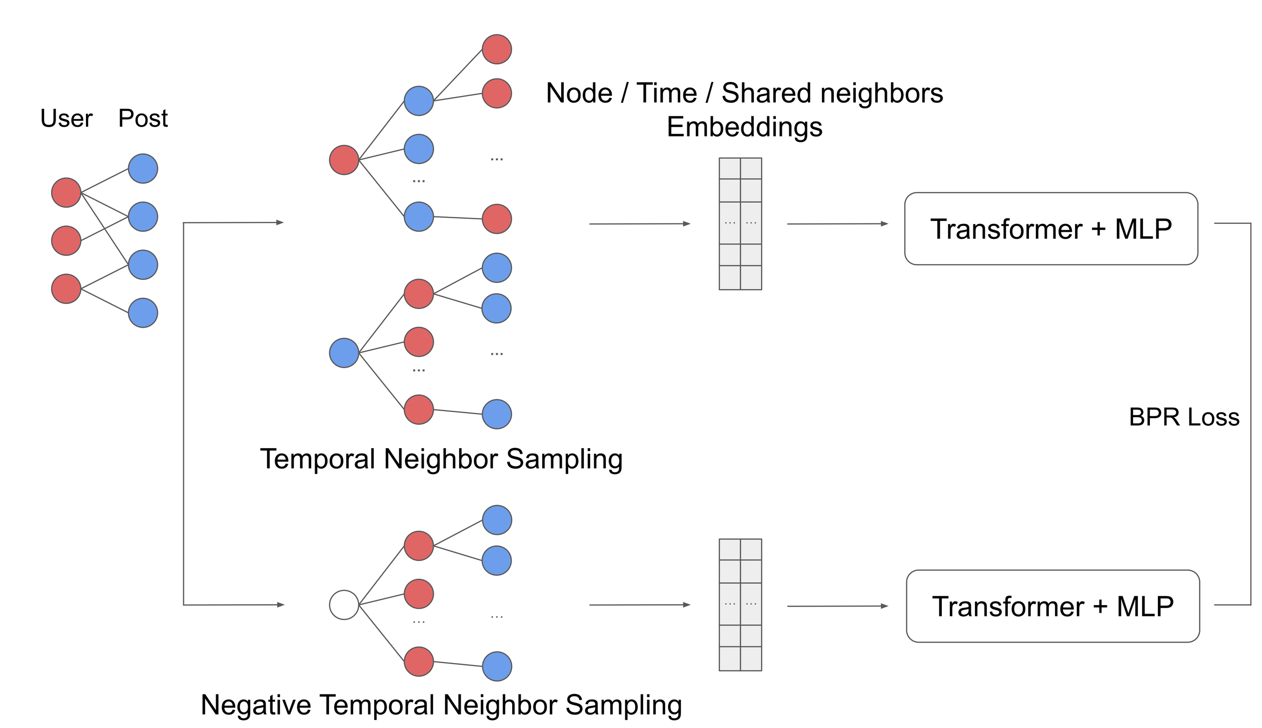

Fig. 4: GraphRec pipeline

- **Temporal Neighbor Sampling:** Take the most recent k user–post interactions. For each positive sample, draw 4 negatives.
- **Embedding Construction:** Combine node, temporal, and shared-neighbor features to form initial embeddings.
- **Transformer + MLP:** Capture sequential/contextual patterns and optimize via BPR.

Results

Key Metrics:

- **Mean Reciprocal Rank (MRR)** – Captures how soon a relevant post appears in the ranked list. The earlier a relevant post appears, the higher the score.
- **Intra-List Diversity (ILD)** – Evaluates how different the recommended items are from one another, ensuring varied content exposure.

Model	MRR ↑	Avg Rank ↓	ILD@10 ↑	Training (1 epoch) ↓	Inference (s/user) ↓
Popularity-Based	0.038	27.25	0.630	NA	**0.019**
MLP-based	0.2 ±	5 ±	0.6 ±	10:48	0.12
GraphRec	**0.271 ±**	**3.69 ±**	**0.704 ±**	17:41	0.17

Discussion

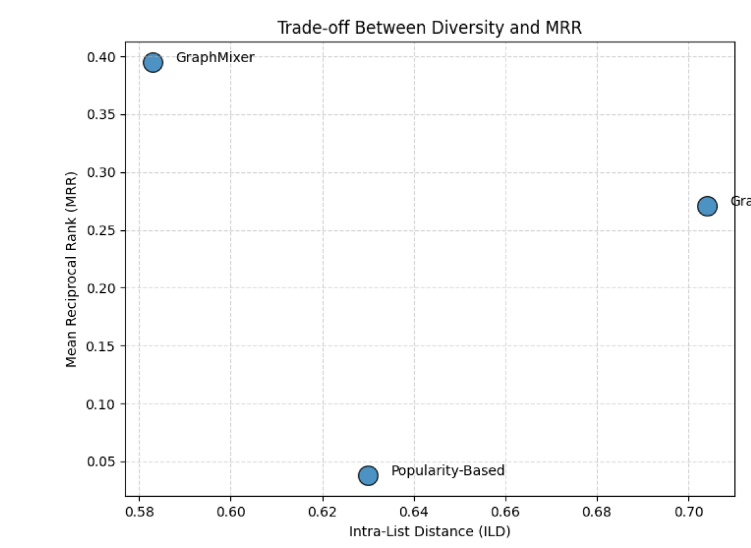

Fig. 5: Trade-off Between Diversity and MRR

Interpretation:

- *GraphRec* has a longer training time but improves diversity and MRR without significantly increasing inference speed. –> need a new plot

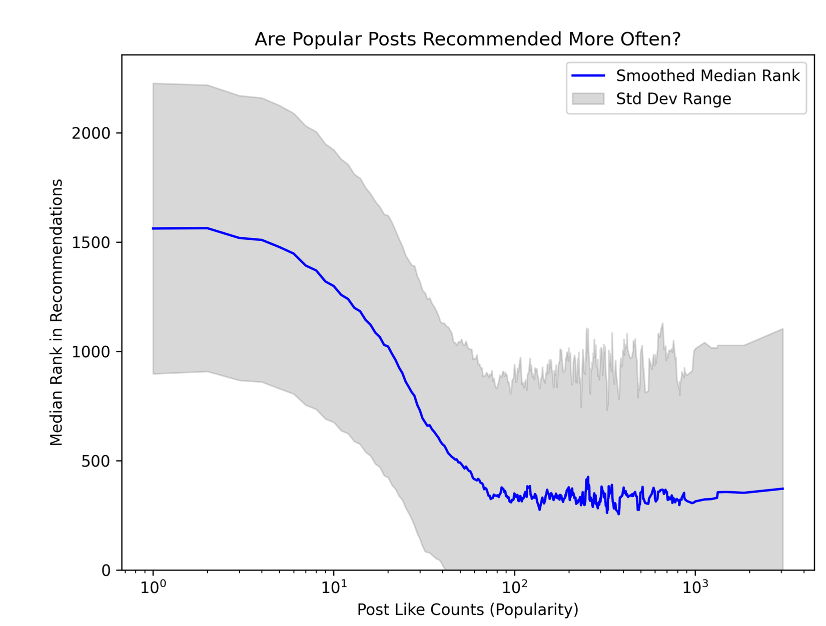

Fig. 6: Median Rank vs. Popularity

- More popular posts are ranked higher on average, but not always at the very top, allowing for some diversity.
- Posts typically need around **50–60 likes** for their median rank to reach **500**, making visibility difficult for less popular posts.
- Adjustments like weighting novelty more or promoting underrepresented content could help balance recommendations.

Acknowledgments

We thank our mentors and the UCSD DSC Capstone Program for their guidance and support. Additionally, we acknowledge the authors of DyGLib [1] for their contributions to dynamic graph learning.

References

[1] Le Yu et al. "Towards Better Dynamic Graph Learning: New Architecture and Unified Library". In: *Advances in Neural Information Processing Systems* (2023).