

Proposal for Implementations in Final Project

DEVIKA SINGH, VARSHA, RINSHI KUMARI
210101036, 210101109, 210108040

1. Individual Fairness

To further generalize the framework, we aim to incorporate individual fairness alongside the group fairness. Individual fairness ensures that users or items with similar characteristics are treated similarly, regardless of their protected attributes, addressing a critical gap in current models. By extending the dynamic fairness learning approach already used to individual fairness, this refinement will allow the system to mitigate bias at a more personalized level, adapting to user-specific needs over time. This ensures a fairer recommendation process for both long-term and individual dimensions.

2. Application Analysis

Analyzing how well the framework from the paper adapts to different recommendation systems is crucial, given its focus on movie recommendations. In e-commerce, where product diversity and seller visibility are key concerns, the method's dynamic fairness approach would need to address the challenge of balancing fairness across rapidly changing inventory and seller performance. Similarly, in point-of-interest (POI) recommendations, factors like location and seasonal popularity shifts make it essential to ensure fair exposure for smaller or lesser-known places. Evaluating the method in these domains would help determine its effectiveness in maintaining fairness in diverse, dynamic environments

3. Enhance Model Interpretability

Introducing techniques to enhance model interpretability, such as LIME (Local Interpretable Model-agnostic Explanations), would be highly applicable in the context of long-term fairness in recommendations. LIME can help identify how specific features influence the recommendations, which is crucial for uncovering biases in the model. By examining feature importance, we can see how attributes like item popularity, user demographics, or other characteristics contribute to the recommendation process. This would allow addressing of any biases, making the model not only fairer but also more transparent. Applying such interpretability techniques would further refine fairness interventions and ensure they align with the model's goals