Meysam Varasteh

meysam.varasteh@colorado.edu | Website | Google Scholar | GitHub | LinkedIn | Boulder, CO

Summary

Ph.D. researcher in machine learning and recommender systems, focusing on making trained recommendation models interpretable and controllable, particularly through post-training concept-based representations and steering mechanisms that enable meaningful interventions while preserving recommendation quality.

Technical Skills

Languages & Tools: Python, C++, PyTorch, TensorFlow, Hugging Face, scikit-learn, CUDA, Docker, Git

ML & Recommender Systems: Transformers, Sequential Recommendation, Ranking, Personalization, Concept Bottleneck Models, Counterfactual Explanations

Selected Research Experience

- **Controllable & Interpretable Sequential Recommender Systems** *Submitted to WWW 2027*

  Developed ConceptRec, a model-agnostic, post-hoc framework that wraps pretrained recommenders (SASRec, LightGCN) in a post-training concept bottleneck, preserving over 97% of nDCG@10 across 3 datasets. Outperformed SAE-based baselines on concept alignment (Overlap@3 up to 0.26 higher) and steering, and introduced a paired steering task and the Push-Pull evaluation metric.
  Steerable and Interpretable Recommendation via Post-Hoc Concept Bottlenecks — **Meysam Varasteh** et al.

- **Provider-Side Transparency in Recommender Systems** *IntRS @ RecSys 2026*

  Developed an interpretable surrogate framework for provider-side transparency that explains system-wide item exposure using structural recommendation features, achieving $R^2 = 0.83$–0.93 across 3 recommender models and 2 datasets and revealing model- and domain-specific exposure drivers.
  Why Didn't More People See It? Recommendation Transparency for Providers — **Meysam Varasteh**, Robin Burke. [paper]

- **Pairwise Counterfactual Explanations for Recommender Systems** *IntRS @ RecSys 2026*

  Developed a comparative counterfactual explanation framework answering **"Why was item A ranked above item B?"** by identifying influential interactions in user histories whose removal reverses the ranking; evaluated across 3 datasets and 2 recommender models, outperforming established explanation baselines in most settings.
  Why This, Not That? Mining User Profiles for Pair-wise Counterfactuals — **Meysam Varasteh**, Veronika Bogina, Noam Koenigstein, Robin Burke. [paper]

- **Fairness Auditing Through the Lens of Counterfactual Explanations** *Spring 2023*

  Developed a counterfactual-based fairness auditing framework to assess whether sensitive groups require unequal effort to achieve favorable model outcomes. Introduced qualification- and effort-based measures and evaluated fairness across multiple classifiers on the Adult Income and Census Income datasets using KL divergence.
  Fairness Auditing Through the Lens of Counterfactual Explanations — Yashar Deldjoo, **Meysam Varasteh**, Yara M. Bahram, Marco Antonio Insabato, Tommaso Di Noia. **[paper]**

- **Review-Aware Hybrid Recommendation** *ICEE 2023*

  Developed a review-aware hybrid recommender system integrating behavioral signals with CNN-based textual representations and attention mechanisms, achieving up to **22.8% improvement** over strong baselines across 3 datasets.
  An Improved Hybrid Recommender System: Integrating Document Context-Based and Behavior-Based Methods — **Meysam Varasteh** et al. [paper]

Education

Ph.D. in Computer Science, University of Colorado Boulder Aug 2023 – Expected May 2028

Advisor: Prof. Robin Burke | Research: Recommender Systems, Interpretability, and Controllable ML

M.Sc. in Electrical and Computer Engineering, University of Tehran 2018 – 2021