

Explainable Uncertainty-Aware Recommendation via Prefix-Tuned Natural Language Generation

Sanjib Kumar Swain

School of Computer Engineering, KIIT University, Bhubaneswar, India

Santosh Kumar Swain

School of Computer Engineering, KIIT University, Bhubaneswar, India

Abstract:

Explainable recommendation systems generate uniformly assertive rationales regardless of upstream uncertainty, eroding user trust when predictions are uncertain. We propose XU-Rec (eXplainable Uncertainty-aware Recommendation), a parameter-efficient architecture that maps uncertainty triples (ue, ua, Δ) to natural language hedging via a 198K-parameter prefix adapter conditioning frozen GPT-2 (117M parameters). The modality-agnostic design accepts uncertainty from any upstream source with architectural non-circularity enforcement (zero text dependency), validated through eight independence tests. We introduce HA-Rich ($H \times R \times C$), a multi-level metric measuring uncertainty-linguistic alignment across lexical, syntactic, and semantic dimensions. Empirical validation on collaborative filtering uncertainty (Amazon Movie Reviews, 100K reviews, 5 seeds) demonstrates learned hedging behavior: $HA_{Rich} = 0.233 \pm 0.030$ with strong correlation ($\rho=0.65$) between interaction-based uncertainty and linguistic hedging, using $650\times$ fewer trainable parameters (198K vs. 128.6M) than full finetuning baselines. The architecture enables future deployment with multimodal perceptual uncertainty (video/audio) without modification.

Keywords:

Explainable recommendation, uncertainty quantification, natural language generation, prefix tuning, hedging appropriateness.