Simplifying Neural News Recommendation: On User Modeling and Training Objectives

Andreea Iana, Goran Glavaš, Heiko Paulheim
Data and Web Science Group, University of Mannheim, Germany
Center for Artificial Intelligence and Data Science, University of Würzburg, Germany
{andreea.iana, heiko.paulheim}@uni-mannheim.de, goran.glavas@uni-wuerzburg.de

Neural News Recommendation Needs Rigorous Evaluation & Simpler Baselines

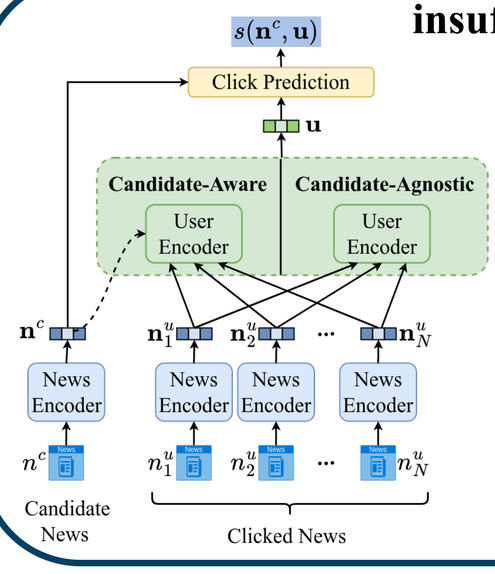
Many factors inhibit fair model comparisons

Usage of proprietary datasets

Often code N/A or hard to use

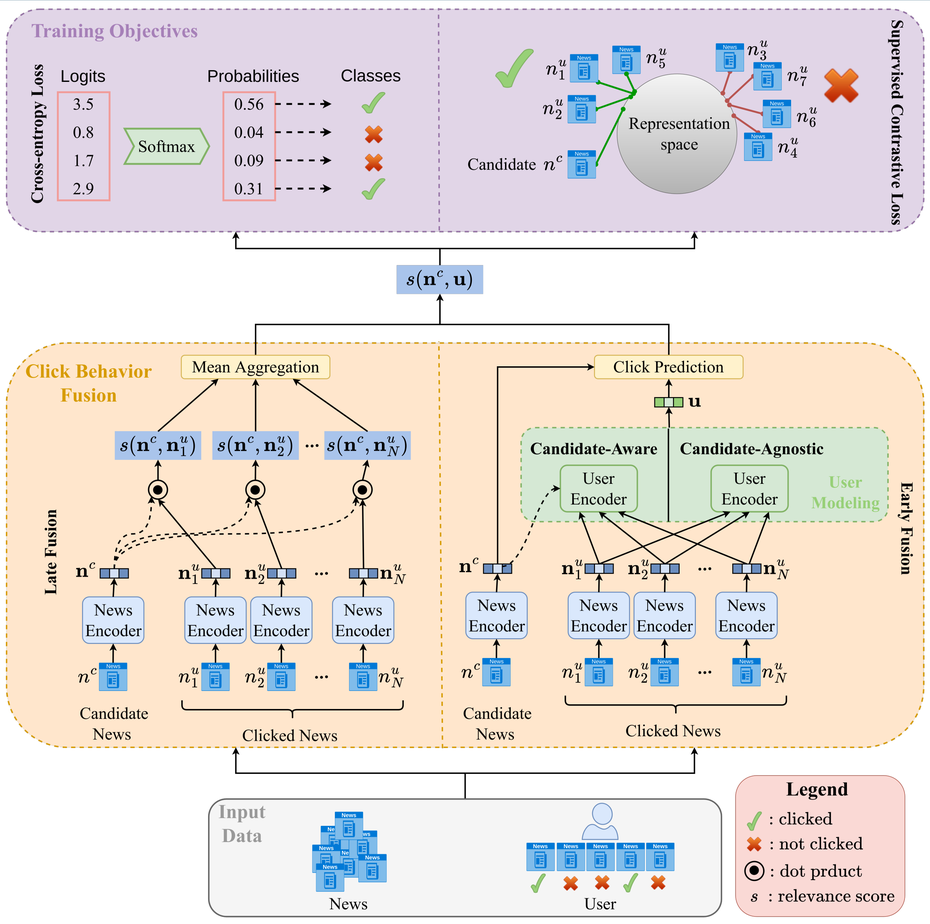
Ad-hoc evaluation protocols

Simpler design & training alternatives insufficiently explored



- Click Behavior Fusion: **early fusion** aggregation of clicked news before comparison w/ candidate
- Training: point-wise classification objective w/ negative sampling

Solution: Unified Evaluation Framework Focused on Key Dimensions



User Modeling

Contextualization against the candidate?

YES **candidate-aware** NO **candidate-agnostic**

Click Behavior Fusion

Explicit user encoder?

YES **early fusion**

aggregate embeddings of clicked news into a user-level representation by means of sequential or attentive encoders

NO **late fusion**

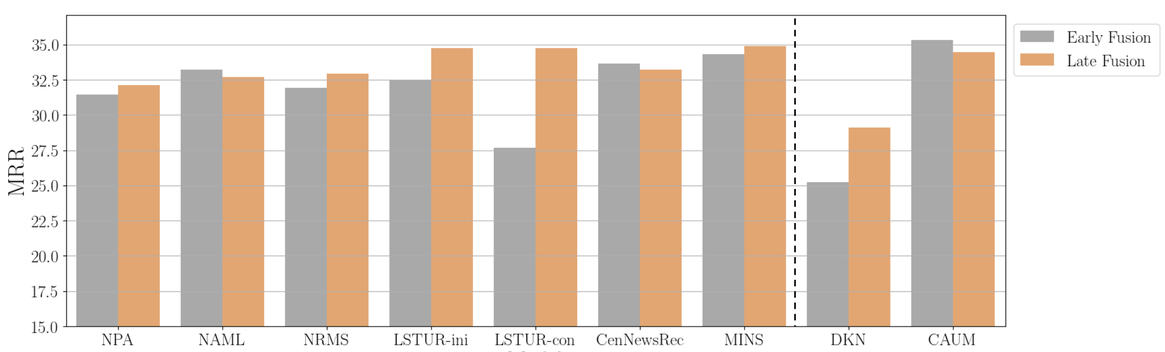
replace user encoders w/ mean-pooling of dot-product scores between embeddings of candidate & clicked news

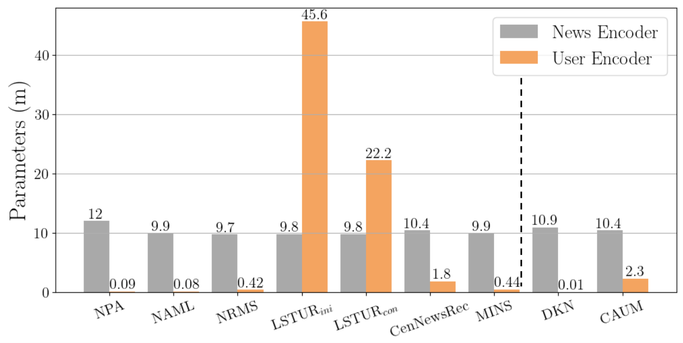
Training Objectives

Point-wise classification objective?

YES **cross-entropy loss** NO **supervised contrastive loss**

Late Fusion Consistently Improves Recommendation Over Complex User Encoders





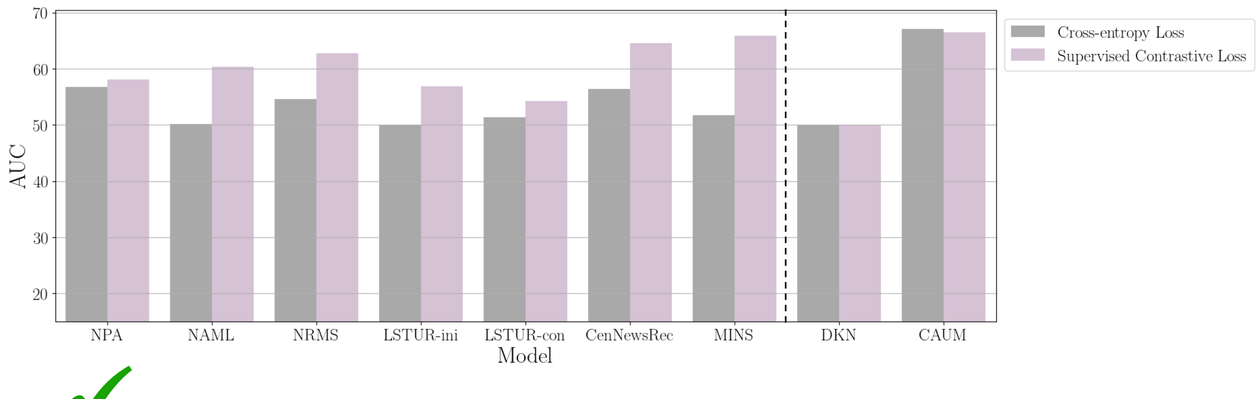
Late fusion

- Consistently on par or better than complex user encoders
- Brings significant reduction in model size

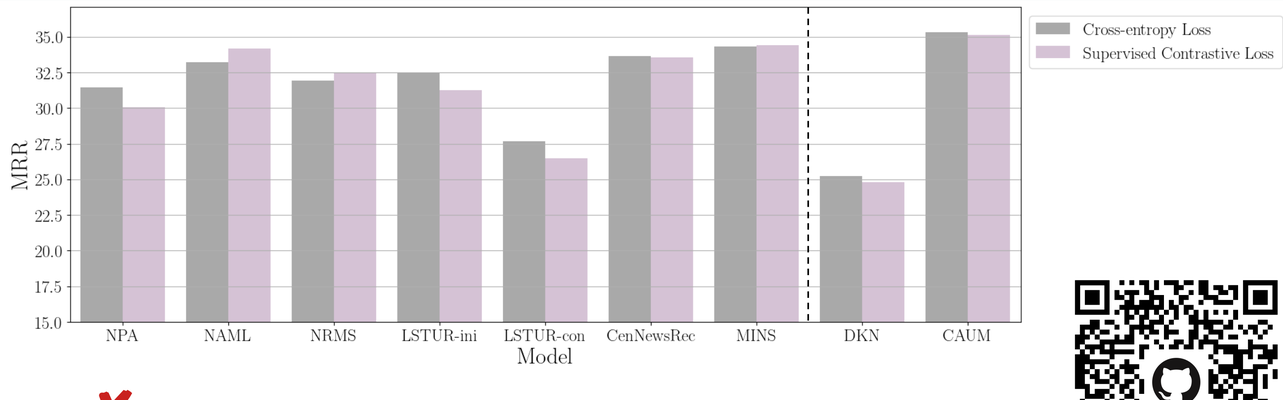
User Modeling

- Complexity of candidate-awareness less beneficial as models perform more similarly under late fusion

Supervised Contrastive Loss Benefits Class Separation



✓ better separation of clicked and non-clicked news



✗ less suitable for distinguishing hard negatives