Neural Methods for Context-aware Product Query Suggestion
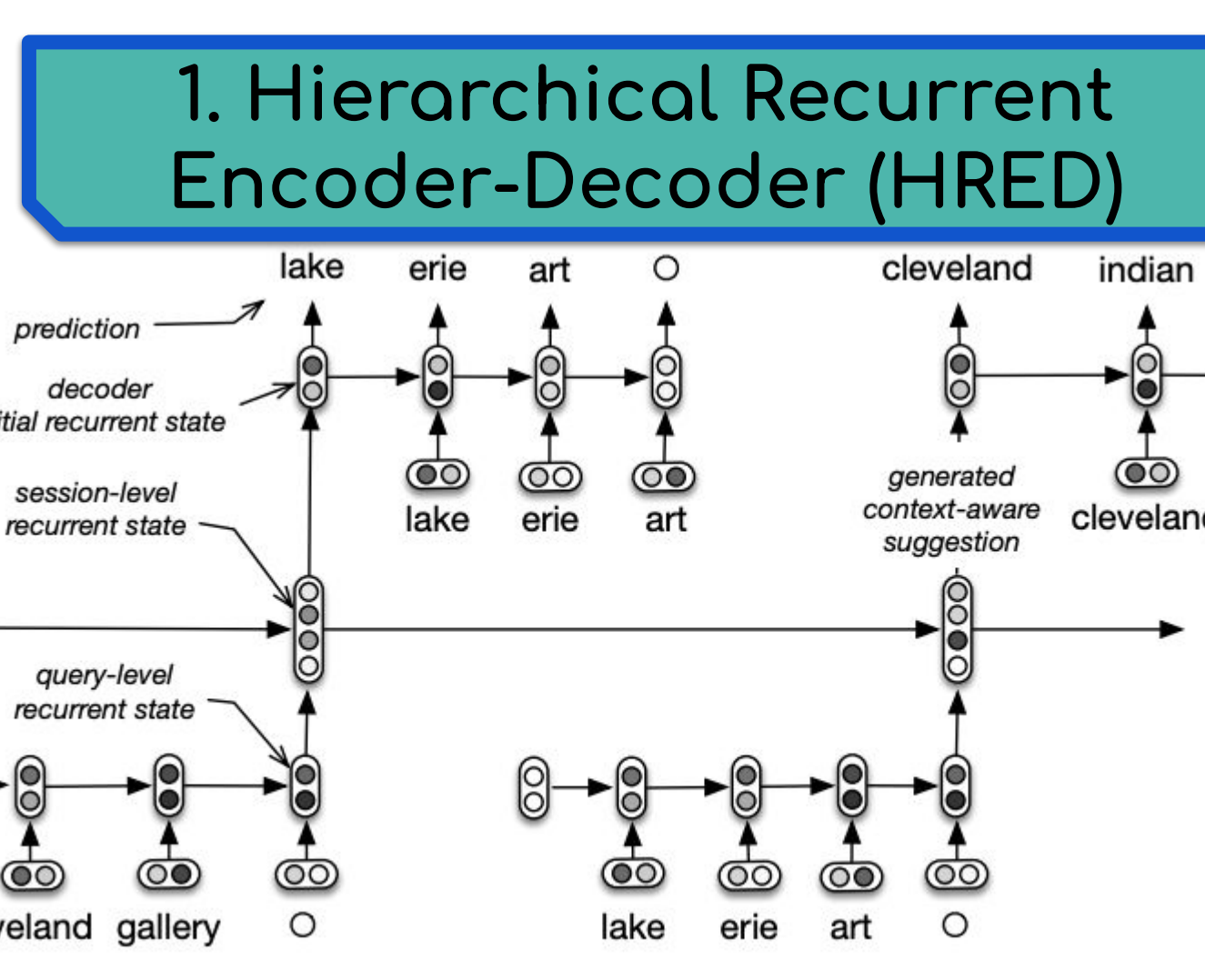
Muthusamy Chelliah

Flipkart, Bangalore, India
muthusamy.c@flipkart.com

ECIR 2022

A. Candidate Generation

1. Hierarchical Recurrent Encoder-Decoder (HRED)
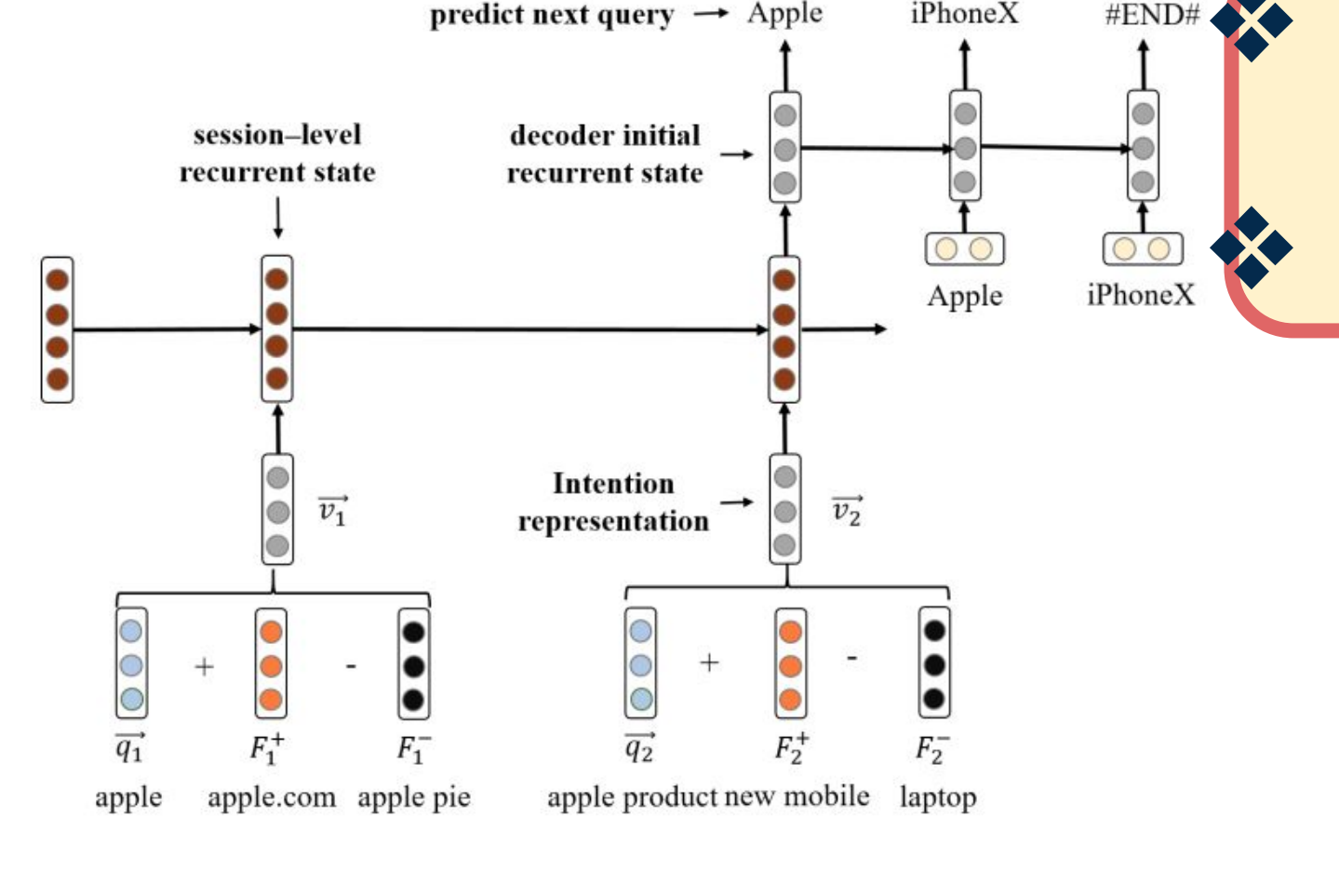Université de Montréal

- ❖ Sparsity of previous queries issued by users
- ❖ Rare/long-tail queries
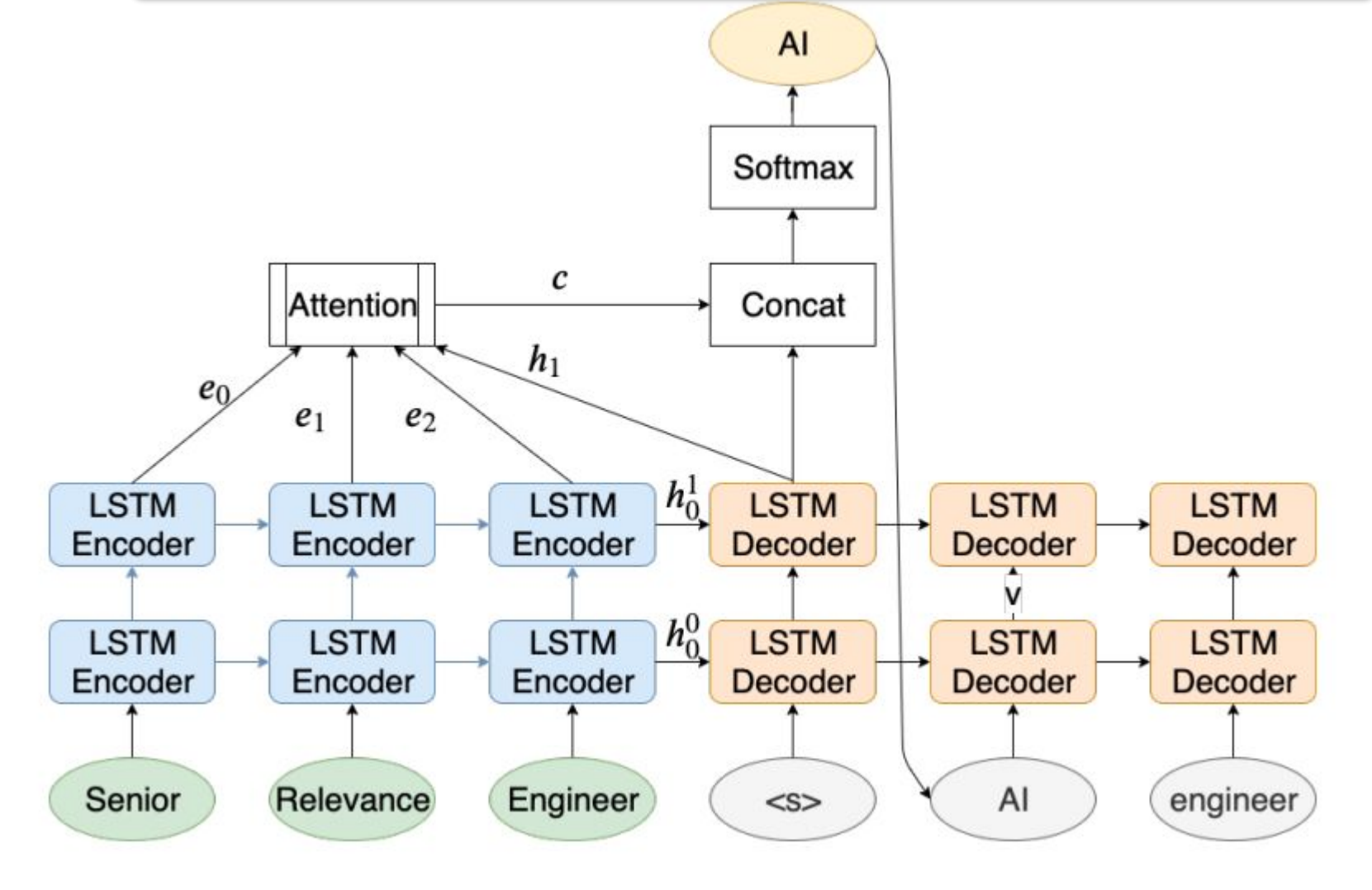- Sequences of arbitrary length in order-sensitive fashion [Sordoni 15]

- ❖ Infrequent/unseen OOV terms retained from earlier queries [Dehghani 17]
- ❖ Query-aware attention for session to capture structure, control scope and detect boundary
- ❖ Decoder to access source words from context

2. Attend-copy-generate (ACG)
Google

B. User interaction

3. Feedback Memory Network (FMN)
清华大学 Tsinghua University

- ❖ Train with history
- ❖ Browse/click converted to attention over top-ranked documents [Wu 18]
- ❖ Combined into memories of query

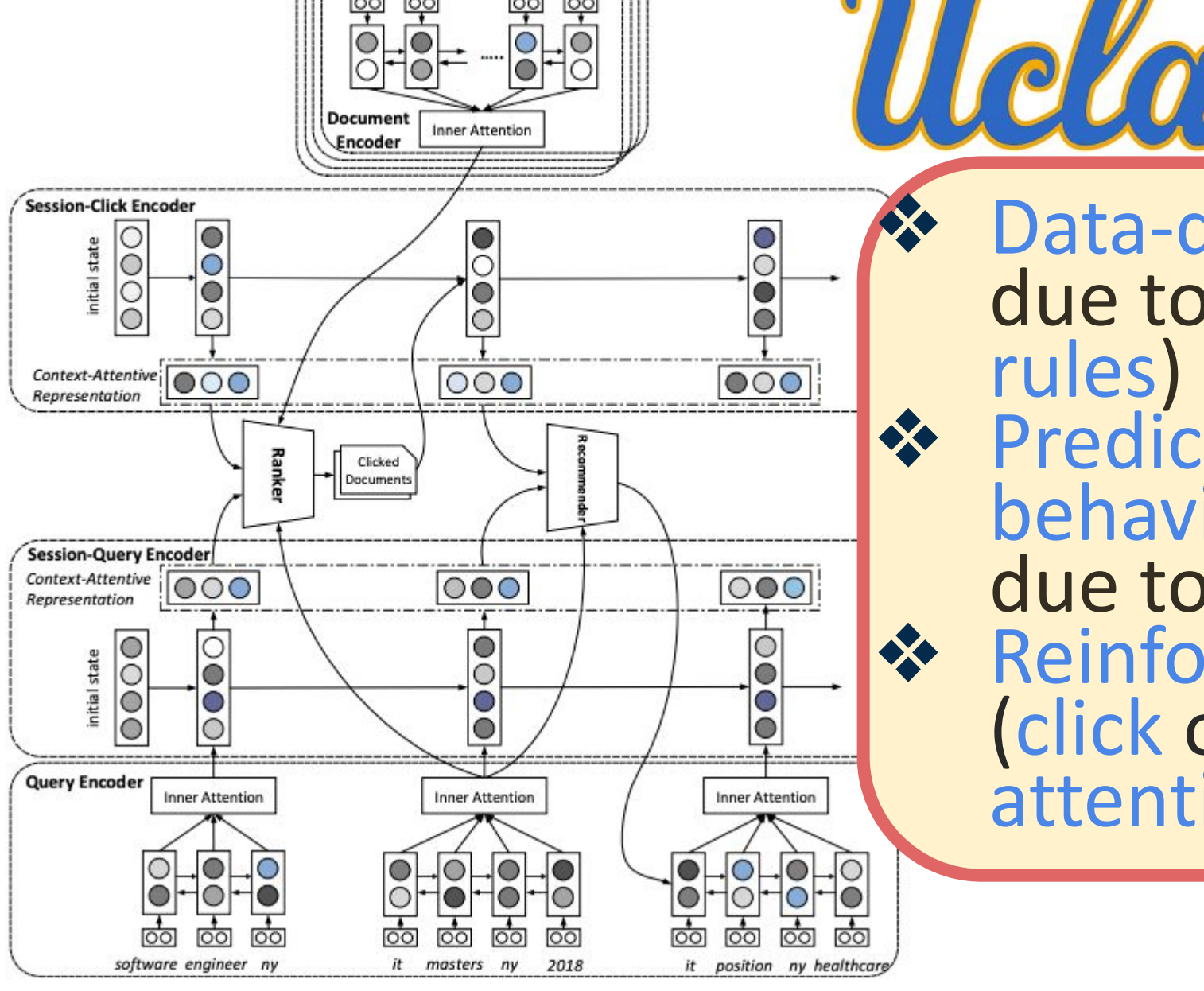### 4. User Feedback Seq2Seq (Ufs2s)
LinkedIn

- ❖ Generation/ranking in the same model [Kazi 20]
- ❖ Feedback is online preference data on machine-generated pairs
- ❖ Separate ranking loss terms (pair-wise penalty) for in-/out-of-domain data types

C. Multi-task Learning
Tencent 腾讯

- ❖ Context embedded in a task helps model intent
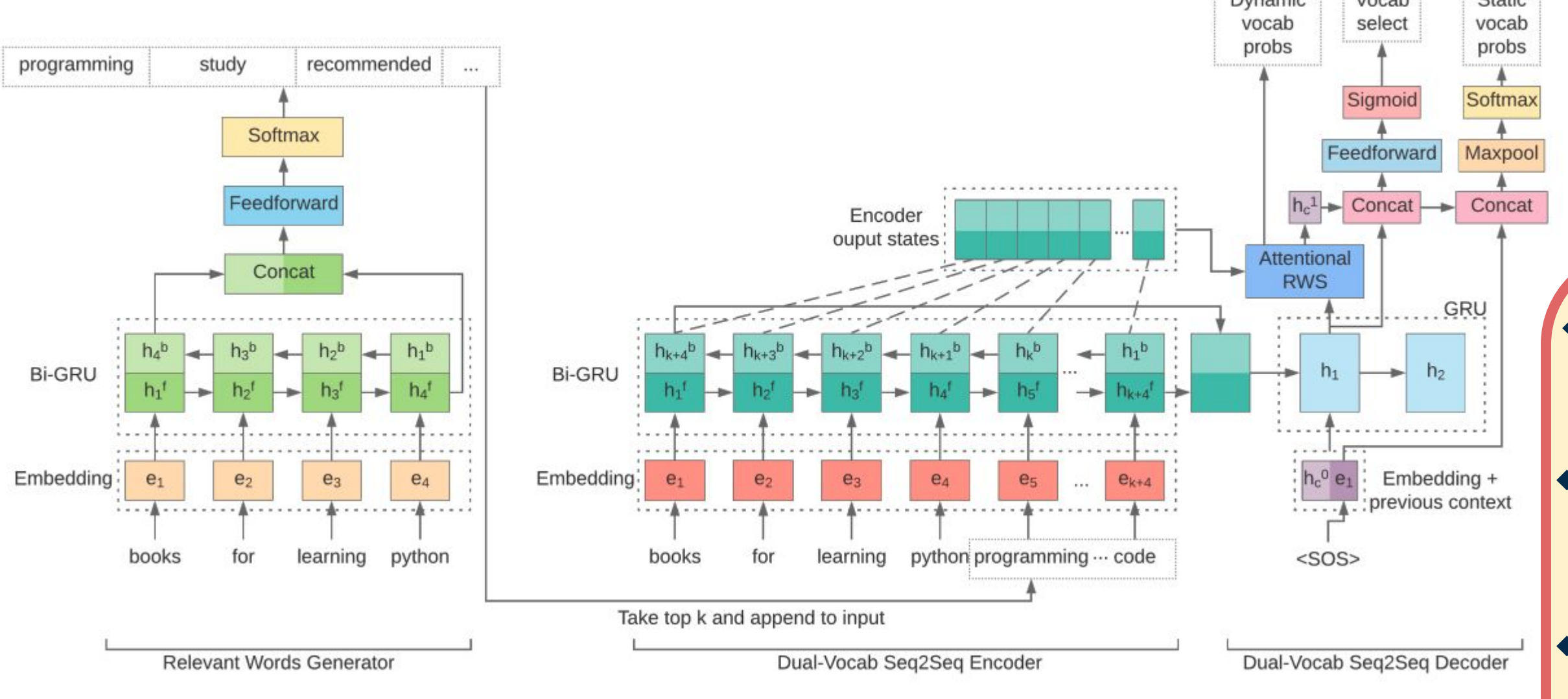- ❖ Leverage domain-specific info. shared across related retrieval tasks

5. Context attentive document ranking (CADR)
Ucla

- ❖ Data-driven solutions not generalizable due to representation form (keywords, rules)
- ❖ Predictive models (HMM) about future behavior forbidden to learn interaction due to predefined space (add term)
- ❖ Reinforcement between retrieval tasks (click on similar query) through context attentive representation [Ahmad 19]

6. Long short-term session search network (LostNet)
University of Amsterdam

- ❖ Learning a user's profile from short-/long-term behavior
- ❖ Long-term memorization, fine-grained profiling [Cheng 21]
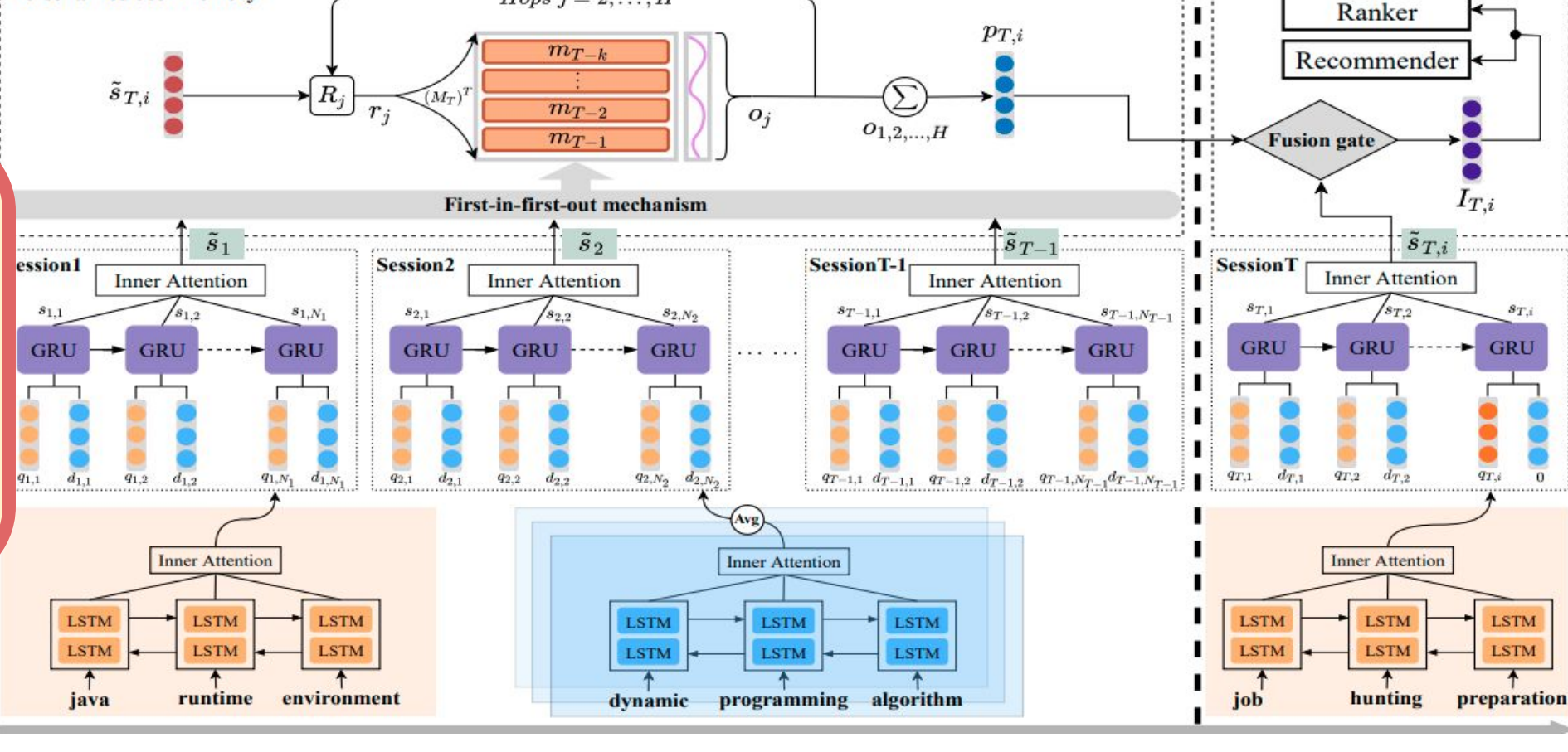
D. Feeds Generation

7. Relevant Word Generator (RWG)

- ❖ Machine translation methods on their own only make suggestions for few well-formed queries
- ❖ User-generated log data: poor readability/content coverage
- ❖ Feed informativeness, coherence

- ❖ Related search recommendation from input query/document title word-by-word [Han 19]
- ❖ Generative Seq2Seq model with a dynamic output vocabulary
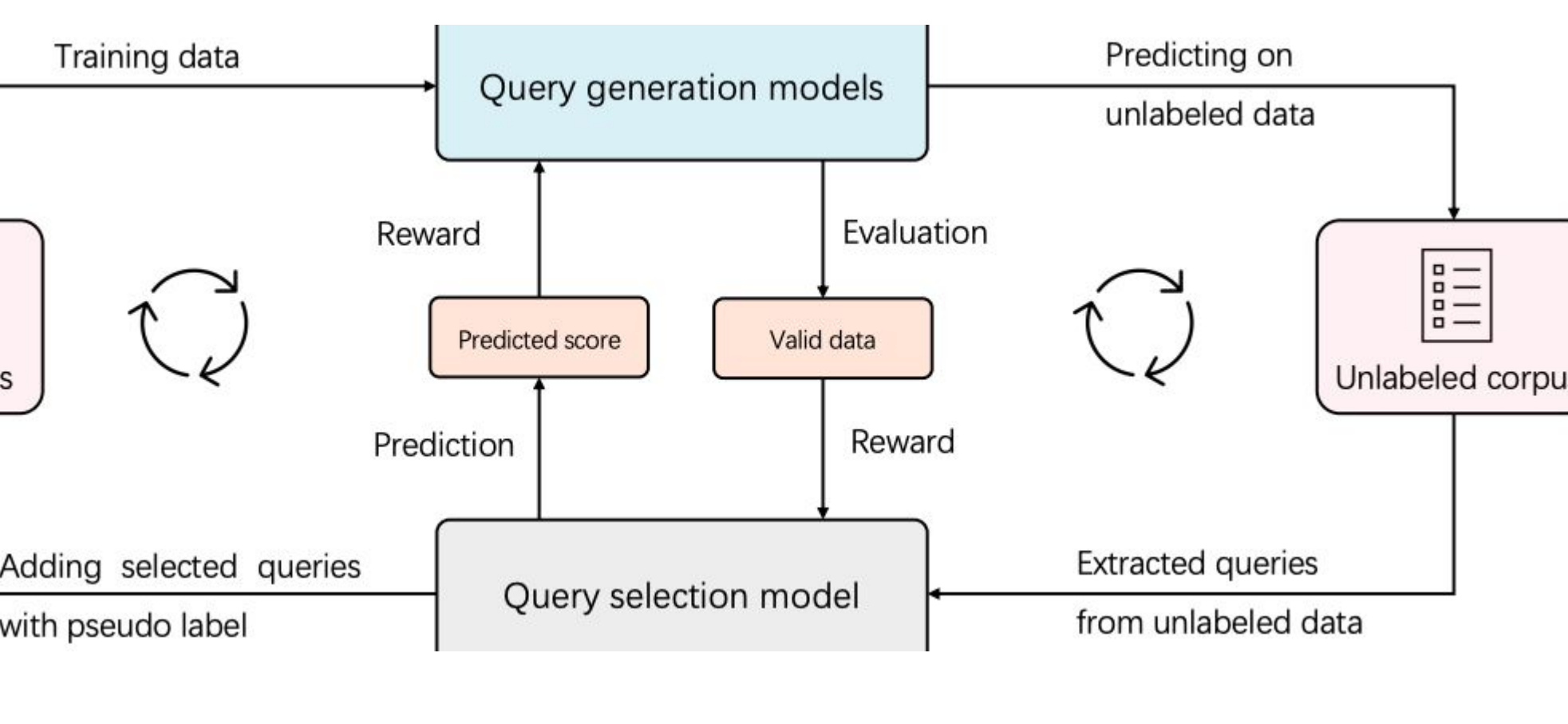- ❖ Relevant context words discovery and context-dependent query generation

8. Title Information Extractor (TitIE)

- ❖ Extract text span (excerpt) from title [Qi 21]
- ❖ Abstractive approach for unsuited titles
- ❖ Improve acceptance by crowdsourced resources
- ❖ Dual/joint learning of generation/selection
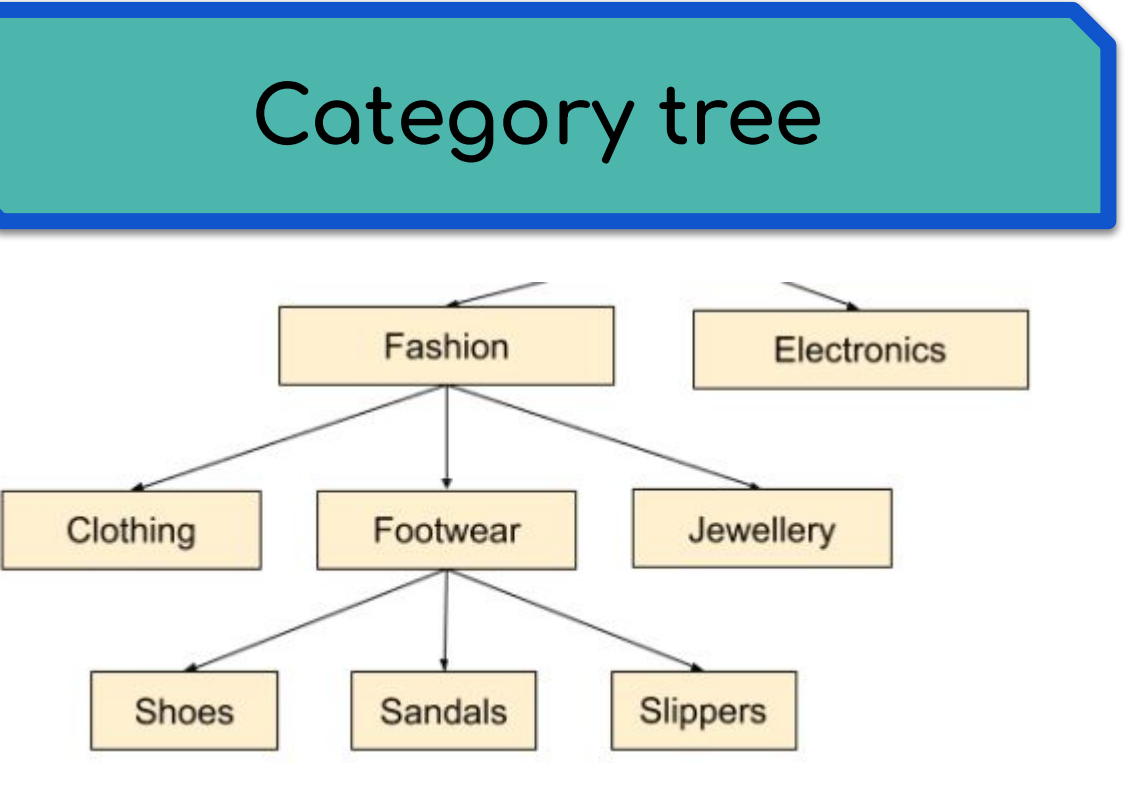 - ➤ predicted data with high scores augments labelled corpus

E. Product query
Flipkart

User session

- ❖ Users follow different shopping patterns
- ❖ Multiple products compared in session
- ❖ Reformulation: log helps learn how queries change [Goyal 21]
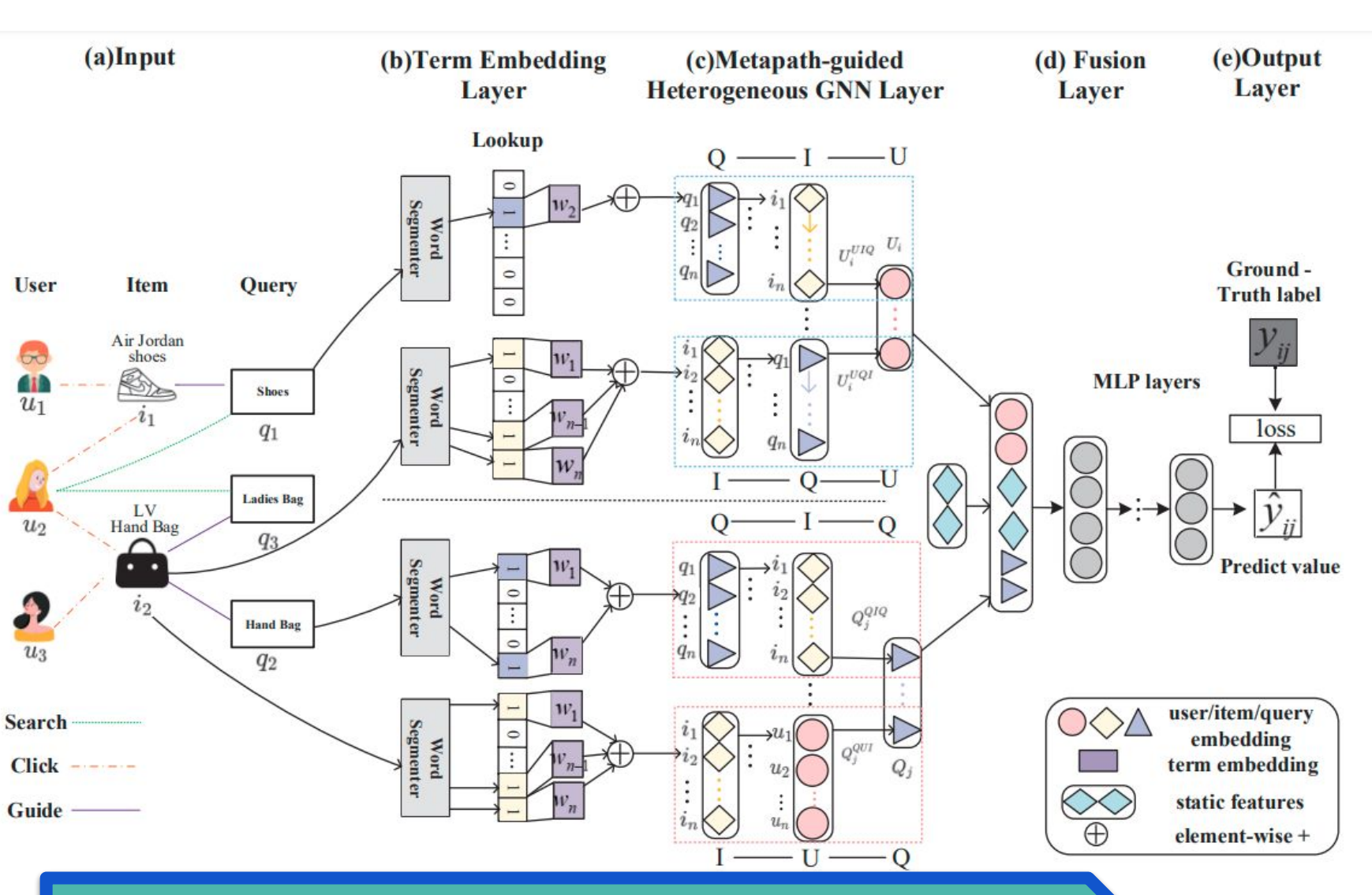- ❖ Success: interaction with results

- ❖ Buyers unfamiliar with category/attributes
- ❖ Understand users' purchase intent
  - ➤ E.g., (shoes,sandals) > (clothing, footwear)
- ❖ Expensive hand-engineered feature sets
- ❖ GBDT ranking based on parameters
  - ➤ Quality/popularity, relevance to prefix
  - ➤ Past actions: cart-add/wishlist/buy
- ❖ Convert aspect into numerical features
  - ➤ Define a function for final scores

Category tree

9. Metapath Guided Network (MPG)
Alibaba.com

- ❖ Interaction among triple- (vs. binary) objects
  - ➤ User, item, query
- ❖ Dynamic intent (query made of words) change
  - ➤ with time/context
  - ➤ Vs. atomic items
- ❖ Interaction relations
  - ➤ Click item, search query, query guide
- ❖ Learn structural feature representation with GNN [Fan 19]
  - ➤ Uniform embeddings with same term space
- ❖ Functions to aggregate neighbor information

F. Context-aware suggestion
- ❖ Query prediction without user input

10. Feedback Interaction NeuralNetwork (FINN)

- ❖ Considering only query/title as signal
  - ➤ Incorrect suggestions per recent intent
  - ➤ *outdoor warm/woman coat jacket*
- ❖ Interaction helps identify noise [Yang 21]
  - ➤ *PF: warm jacket, women sneaker, long skirt*
  - ➤ *Negative Feedback: laptop->MacBook Air*
- ❖ Filter attention (FAT)
  - ➤ Future Negative Feedback to filter PF
  - ➤ *Outdoor warm jacket -> warm jacket*

References

[Sordoni 15] A Hierarchical Recurrent Encoder-Decoder for Generative Context-Aware Query Suggestion, CIKM
[Dehghani 17] Learning to Attend, Copy, and Generate for Session-Based Query *Suggestion*, CIKM
[Jiang 18] RIN: Reformulation Inference Network for Context-Aware Query Suggestion, CIKM
[Wu 18] Query Suggestion with Feedback Memory Network, WebConf
[Li 19] Click feedback-aware query recommendation using adversarial examples, WebConf
[Kazi 20] Incorporating User Feedback into Sequence to Sequence Model Training, CIKM Applied Science
[Ahmad 19] Context Attentive Document Ranking and Query Suggestion, SIGIR
[Cheng 21] Long Short-Term Session Search: Joint Personalized Reranking and Next Query Prediction, WebConf
[Han 19] A Deep Generative Approach to Search Extrapolation and Recommendation, KDD
[Qi 21] Dual Learning for Query Generation and Query Selection in Query Feeds Recommendation, CIKM Applied Science
[Gollapudi 11] Shopping for Products You Don't Know You Need, WSDM
[Hasan 11] Query Suggestion for E-Commerce Sites, WSDM

References

[Sanjanwala 21] https://blog.flipkart.tech/predicting-your-next-query-even-before-you-type-83487a34109d
[Goyal 21] https://tech.flipkart.com/building-personalized-autosuggestion-9e705d5bf5f8
[Maji 19] Addressing vocabulary gap in e-commerce search, SIGIR short
[Maji 20] A Regularised Intent Model for Discovering Multiple Intents in E-Commerce Tail Queries, ECIR
[Fan 19] Metapath-guided Heterogeneous Graph Neural Network for Intent Recommendation, KDD
[Yang 21] FINN: Feedback Interactive Neural Network for Intent Recommendation, WebConf
[Yin 20] Learning to Generate Personalized Query Auto-Completions via a Multi-View Multi-Task Atten Approach, KDD Applied Data Science
[Chen 17] Personalized Query Suggestion Diversification, SIGIR short
[Su 21] Modeling Intent Graph for Search Result Diversification, SIGIR
[Yan 21] Diversification-Aware Learning to Rank using Distributed Representation, WebConf