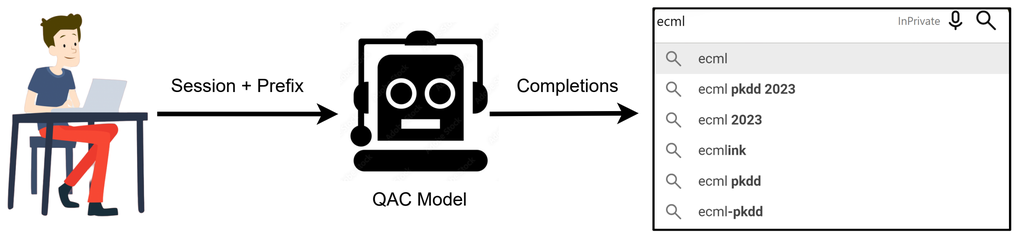
TRIE-NLG: Trie Context Augmentation to Improve Personalized Query Auto-Completion for Short and Unseen Prefixes

Kaushal Kumar Maurya[1] **and Maunendra Sankar Desarkar**[1] **and Manish Gupta**[2] **and Puneet Agrawal**[2]

[1]Department of CSE, IIT Hyderabad, India
[2]Bing AutoSuggest, Microsoft, Hyderabad, India

Email: cs18resch1103@iith.ac.in

1 Background



- *Query Auto Completion (QAC)* recommends a list of relevant complete queries for the partially typed search query (i.e., prefix)
- Helps in: (1) Saving keystrokes (2) Understanding the user's search intent and (3) Assisting in formulating complex search queries

2 Knowledge Gap

1. **Problem of Unseen and Short Prefixes (Based on Stats from Bing Query Log):**
   - Approximately 12% prefixes do not have any completions from Trie
   - Approximately 44% prefixes have a length of less than 6 characters
2. **Trie/Ranking Models:**
   - (+) Suggestions are more meaningful as they come from user log
   - (-) No personalization
   - (-) Provide a limited number of suggestions
   - (-) No suggestions for *Unseen Prefixes*
3. **NLG Models:**
   - (+) Can model personalization
   - (+) Generate suggestions for unseen prefixes
   - (-) For *Short Prefixes*, suggestions are bad due to limited context

3 Research Question

Factors to be considered in the Modeling: (1) Personalization (2) Trie Context and (3) Deployable Latency

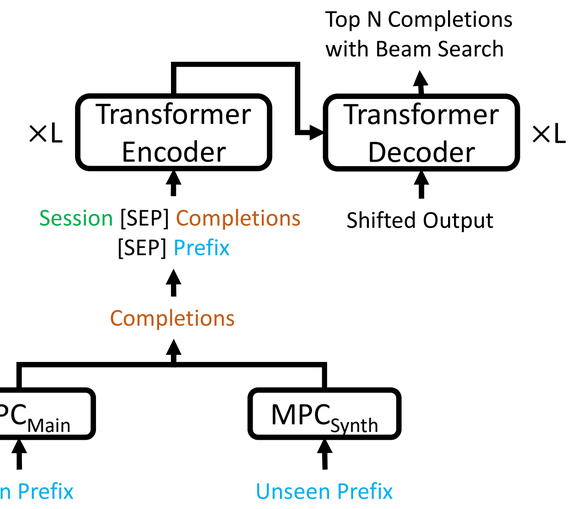
Given current session $s = \{sq_1, sq_2 \dots sq_n\}$, prefix p and external trie context $e = \{c_1, c_2, \dots c_m\}$, generate top-$N$ query completions $q = \{q_1, q_2 \dots q_N\}$ conditioned on s, e and p.

4 Methodology: TRIE-NLG Model

- **MPC_{Main}:** Trie context extraction for seen prefixes
- **MPC_{Synth}:** Synthetic context extraction for unseen prefixes [1], come from suffixes of past queries:

A Session Query: University of west florida
Candidate Suffixes: (1) florida (2) west florida and (3) of west florida
MPC_{Synth}: { *University of west* : florida, *University of* : west florida, *University* : of west florida}

- **Context Augmentation:** MPC_{Main} and MPC_{Synth} context augmentation in Pre-trained Language Model (PLM) based NLG model.



5 Results: Unseen and Short Prefixes

Bing Unseen Dataset			
Models	**ΔMRR**	**$\Delta BLEU_{RR}$**	**ΔBLEU**
MPC_{Train}	-	-	-
MPC_{Main}	-	-	-
MPC_{Main} + MPC_{Synth}	0.00	0.00	0.00
GRM	-	-	-
Seq2Seq LSTM	5.34	16.61	91.45
Seq2Seq Transformer	10.53	24.20	99.29
T5	21.60	27.33	103.98
BART	34.53	29.03	110.35
BART + ITC	33.05	29.12	111.00
BART + MPC_{Main}	34.10	28.04	110.80
TRIE-NLG	**59.74**	**33.02**	**123.07**

Results for Short Prefixes (Bing)			
Models	**ΔMRR**	**$\Delta BLEU_{RR}$**	**ΔBLEU**
Prefix Length in [1-5]			
BART	21.6	-17.6	2.9
BART + MPC_{Main}	52.1	-16.8	10.1
TRIE-NLG	**53.2**	**-15.9**	**11.4**
Prefix Length in [6-10]			
BART	38.5	38.4	56.8
BART + MPC_{Main}	55.2	49.3	72.1
TRIE-NLG	**56.4**	**50.1**	**73.6**
Prefix Length 10+			
BART	49.1	190.0	120.1
BART + MPC_{Main}	55.6	218.5	145.1
TRIE-NLG	**60.7**	**221.1**	**149.5**

Sample Generations from TRIE-NLG

Example-1: Seen Short Prefix		Example-2: Unseen Prefix	
Session: kysportsradio || kysportsradio || cincinnati reds || cincinnati reds || espn sports || espn sports || ebth || ebth.com || ebth.com || cnn news || cnn news || politico news Prefix: p Correct Query: politico		Session: hurricane resistant || hurricane lines || houston crap || houston crap plan || hurricane climate Prefix: houston climate actio Correct Query: houston climate action plan	
Completions (MPC_{Main}): 1. pinterest 2. paypal 3. pittsburgh penguins 4. pandora 5. prime video 6. paypal login account 7. pennlive 8. pogo official site	Completions (Trie-NLG): 1. politico 2. profootballtalk 3. politico news 4. pittsburgh pirates 5. pogo official site 6. page tour 7. philadelphia inquirer 8. pennlive	Completions(MPC_{Main}): None Completions(MPC_{Synth}): 1. houston climate action policy	Suggestions (Trie-NLG): 1. houston climate action plan 2. houston climate action policy 3. houston climate action play 4. houston climate plan action 5. houston climate action plan tx 6. houston climate action program 7. houston climate action plan plan 8. houstonclimate action plan

6 Conclusions

- We motivated the need of incorporating trie and session context to improve auto-completions for short and unseen prefixes.
- We proposed a novel architecture, TRIE-NLG; this is the first attempt of *Trie knowledge augmentation* in NLG models for personalized QAC.

Reference and Acknowledgement

[1] Bhaskar Mitra and Nick Craswell. 2015. Query auto-completion for rare prefixes. In CIKM 2015.

- This research and conference travel is partially funded by Microsoft Academic Partnership Grant 2022.