Knowledge-Enriched Natural Language Generation

EMNLP 2021 Tutorial

Wenhao Yu¹, Meng Jiang¹, Zhiting Hu², Qingyun Wang³, Heng Ji³, Nazneen Rajani⁴

¹ University of Notre Dame  ² University of California San Diego  
³ University of Illinois at Urbana-Champaign  ⁴ Salesforce Research
This part: General principles and methodologies for integrating knowledge into NLG

Overview:

• Knowledge-enhanced model architectures
  • Attention/copy mechanisms
  • Graph neural models

• Knowledge-enhanced learning
  • Auxiliary loss/tasks
  • Reinforcement learning with knowledge-informed rewards
  • Learning with knowledge constraints

• Knowledge-enhanced inference
  • Steered decoding
  • Prompts
Knowledge-enhanced inference

- Integrate knowledge during the text decoding process
- Can be applied to pretrained language models (e.g., GPT-2/3, T5) for knowledge-enhanced NLG
Inference (I): Steered decoding

• Guide the decoding by changing the generation distribution
• Controlling LMs with another (relatively smaller) model trained for desired attributes
• Examples: PPLM, GeDi, DeLorean, DExperts, FUDGE, MoFE
Inference (I): Steered decoding

- Use a single-layer discriminative classifier as the guide

Dathathri et al., “Plug and Play Language Models (PPLM)”
Inference (I): Steered decoding

- Use another LM as a generative discriminator to guide decoding
- Examples: GeDi (Krause et al., ‘20) and FUDGE (Yang et al., ‘21)

\[ P(\text{pos}|x_{1:t}) = \frac{P(x_{1:t}|\text{pos})}{P(x_{1:t}|\text{neg}) + P(x_{1:t}|\text{pos})} \]

Krause et al., “Generative Discriminator (GeDi) guided decoding”
Inference (I): Steered decoding

- Use two LMs to constrain decoding

Liu et al., “DExperts”
Inference (I): Steered decoding

- Using RL to constrain decoding for factual consistency

Choubey et al., “Mixture of Factual Experts (MoFE)”
Inference (I): Steered decoding

- Using text to constrain decoding

Qin et al., “DELOREAN”
Inference (II): Prompts

- Guide the decoding by changing the generation distribution
- Low parameter alternative to finetuning LMs
Inference (II): Prompts

3 broad directions to infer knowledge using prompts:

1. zero-shot, eg: “The capital of Canada is [MASK]”

2. few-shot, eg: “The capital of France is Paris. The capital of Canada is [MASK]”

3. with additional context, eg: “Ottawa sits on the Ottawa River at the border between Quebec and Ontario. The capital of Canada is [MASK].”
Inference (II): Prompts

Zero-shot setting

Petroni et al., “LMs as KBs”
Inference (II): Prompts

Few-shot setting/ via demonstration

Gao et al., 2021
Inference (II): Prompts

With additional context

Bian et al., “Knowledge enhanced Commonsense QA”
Conclusion

**This part**: General principles and methodologies for integrating knowledge into NLG

- Knowledge-enhanced *model architectures*
- Knowledge-enhanced *learning*
- Knowledge-enhanced *inference*
  - Steered decoding
  - Prompts