EMNLP 2021 Tutorial

Knowledge-Enriched Natural Language Generation

Wenhao Yu¹, Meng Jiang¹, Zhiting Hu², Qingyun Wang³, Heng Ji³,⁴, Nazneen Rajani⁵

¹ University of Notre Dame  ² University of California San Diego  ³ University of Illinois at Urbana-Champaign  ⁴ Amazon Scholar  ⁵ Salesforce Research
<table>
<thead>
<tr>
<th>Time (AST / EST)</th>
<th>Content</th>
<th>Tutor</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00 am – 09:10 am (08:00 am – 08:10 am)</td>
<td>Motivation and Key Questions</td>
<td>Meng Jiang</td>
</tr>
<tr>
<td>09:10 am – 09:50 am (08:10 am – 08:50 am)</td>
<td>General Methods of Integrating Knowledge into NLG</td>
<td>Zhititing Hu, Nazneen Rajani</td>
</tr>
<tr>
<td>09:50 am – 10:15 am (08:50 am – 09:15 am)</td>
<td>NLG Methods Enhanced by Various Knowledge Source</td>
<td>Wenhao Yu</td>
</tr>
<tr>
<td>10:15 am – 10:45 am (09:15 am – 09:45 am)</td>
<td>Coffee break</td>
<td></td>
</tr>
<tr>
<td>10:45 am – 11:10 am (09:45 am – 10:10 am)</td>
<td>NLG Methods Enhanced by Various Knowledge Source</td>
<td>(cont.)</td>
</tr>
<tr>
<td>11:10 am – 12:10 pm (10:10 am – 11:10 am)</td>
<td>Applications, Benchmark Datasets, Code Library</td>
<td>Qingyun Wang</td>
</tr>
<tr>
<td>12:10 pm – 12:30 pm (11:10 am – 11:30 am)</td>
<td>Remaining Challenges and Future Directions</td>
<td>Heng Ji</td>
</tr>
</tbody>
</table>
Text generation is needed everywhere in our life!

- Leveraging machine intelligence to make many things EASY and FAST

What should I wear outside today?

The temperature will be 85 at noon. You can wear short sleeve shorts.

Dialog systems

Summarization

"... Her husband was one of 17 people killed in January’s terror attacks in Paris. ... Valerie Braham said to the assembled crowd. ..."

"... Philippe Braham was killed in the January’s terror attacks. ..."

Machine Translation

Paraphrasing

"... The game ended with the umpire making a bad call, and if the call had gone the other way, the Blue Whales might have actually won the game. It wasn't a victory, but I say the Blue Whales look like they have a shot at the championship, especially if they continue to improve."

"... The match ended with the referee calling wrongly, and if the call went the other way, the Blue Whales could already win the match. It was not a victory, but I say that the Blue Whales seem to have a chance in the championship, especially if they keep improving."

"... (Two years after, Maria was born ... ) ... Her brother gave her a hug. ...

"... 她的哥哥给了她一个大大的拥抱. ..."
Text generation is needed everywhere in our life!

Image Captioning: Image-to-Text Generation

Sports News Generation: Table-to-Text Generation

Not the main focus of this tutorial.
Text-to-text generation is needed everywhere!

\[
P(Y|X) = P(y_1, \ldots, y_m|x_1, \ldots, x_n) = \prod_{t=1}^{m} p(y_t|X, y_1, \ldots, y_{t-1}), \text{ when } Y \text{ is text and } X \text{ is text.}
\]

- What should I wear outside today?
  - The temperature will be 85 at noon. You can wear short sleeve shorts.

**Dialog systems**

- "... Her husband was one of 17 people killed in January’s terror attacks in Paris. ... Valerie Braham said to the assembled crowd. ..."

**Summarization**

- "... Philippe Braham was killed in the January’s terror attacks. ..."

**Machine Translation**

- "... The game ended with the umpire making a bad call, and if the call had gone the other way, the Blue Whales might have actually won the game. It wasn’t a victory, but I say the Blue Whales look like they have a shot at the championship, especially if they continue to improve."

**Paraphrasing**

- "... The match ended with the referee calling wrongly, and if the call went the other way, the Blue Whales could already win the match. It was not a victory, but I say that the Blue Whales seem to have a chance in the championship, especially if they keep improving."
**Knowledge** is needed everywhere in **text generation!**

- Mistakes may occur due to lack of knowledge.

**What should I wear outside today?**

*The temperature will be 65°F at noon. You can wear short sleeve shorts.*

**Dialog systems**

*... Her husband was one of 17 people killed in January’s terror attacks in Paris. ... Valerie Braham said to the assembled crowd. ..."

*... Valerie Braham
Philippe Braham was killed in the January’s terror attacks. ..."

**Summarization**

*... The game ended with the umpire making a bad call, and if the call had gone the other way, the Blue Whales might have actually won the game. It wasn’t a victory, but I say the Blue Whales look like they have a shot at the championship, especially if they continue to improve."

*... The match ended with the referee calling wrongly, and if the call went the other way, the Blue Whales could never already win the match. It was not a victory, but I say that the Blue Whales seem to have a chance in the championship, especially if they keep improving."*

**Machine Translation**

*... (Two years after, Maria was born ...) ... Her brother gave her a hug. ..."

*... 她的弟弟哥哥给了她一个大大的拥抱. ...”*

**Paraphrasing**

*... (Two years after, Maria was born ...) ... Her brother gave her a hug. ...”*
Knowledge-Enriched Natural Language Generation (KE-NLG)

Key questions: (Outline of this tutorial)
- Where does the Knowledge come from? In other words, what are the Knowledge sources?
- What are the data/representations of Knowledge?
- What are the methods for integrating Knowledge into NLG models? How to utilize different types of Knowledge?
- What are the concrete NLG models/techniques that have been enhanced by Knowledge? What are their advantages and disadvantages?
- Where can we find benchmark datasets, code library, and hands-on tutorial for doing research on this topic?
- What are the remaining challenges and future directions?
Material for this tutorial

• Website: https://kenlg-tutorial.github.io/
  • Schedule
  • Slides
  • Tutor information

• A pre-print survey on the topic of Knowledge-enhanced NLG:

• A reading list and a collection of open-source code:
  • https://github.com/wyu97/KENLG-Reading

Enjoy