Assignment 2 comments

Jindřich Libovický and Jindřich Helcl, 9.5. 2024
Assignment 2 - translating into English

● Prompt matters
  ○ Sometimes including weird stuff like quadruple colons also helps.
  ○ Shorter prompts tend to be better
  ○ Specifying the languages might be a good idea

● Generally, greedy decoding performs better than sampling
  ○ Beam search might be best

● Different models work differently for different languages
  ○ Tower often the best
  ○ Aya generates short outputs - it was trained on sentence-split data!

● Few shot prompting did not seem to help much
  ○ unless you provide the right few shots (sentences from the same article for example)

“Translation: I have no idea what he is talking about” – Mistral, 2024
Assignment 2 - translating into X

- Translations to non-English languages often end up English
  - Maybe translate the prompt to the target lang?

- When pivoting, you can use different models

- Translating into lower-resource languages often fail completely
  - (Georgian, Macedonian, Persian, Slovak, Ukrainian, ...)

Reading Papers

Jindřich Libovický and Jindřich Helcl, 9.5. 2024
Assigned papers


Time to vote!

https://forms.gle/JFdf1BmzbPnqAd9LA
# Results

<table>
<thead>
<tr>
<th>Paper</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blind Spots and Blinds</td>
<td>0</td>
</tr>
<tr>
<td>Toolformer</td>
<td>6</td>
</tr>
<tr>
<td>Leak, Cheat, Repeat</td>
<td>4</td>
</tr>
<tr>
<td>Mistral 7B</td>
<td>5</td>
</tr>
<tr>
<td>LangBridge</td>
<td>3</td>
</tr>
</tbody>
</table>
Some Leads on Paper Quality
Others tips

- Put numbers into context:
  Leaderboards for datasets [https://paperswithcode.com](https://paperswithcode.com)

- Check social networks (papers are often discussed on Twitter/X)

- What is the track record of the lab that published it?

- Was it published? Where? (CORE Conference Ranking)
  Some venue provide reviews

- How many times was it cited? (Relevant only after some time.)

- Is the code available? (If yes, authors did not make up the results.)
<table>
<thead>
<tr>
<th>Metadatas for papers under consideration</th>
<th>Was it published?</th>
<th>Reviews available</th>
<th>Cited</th>
<th>How long it is around?</th>
<th>Code available</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blind spots and blinds</td>
<td>No, probably rejected</td>
<td>No (but it was in ARR in Oct 2023)</td>
<td>$4 \times$ 1.5 times a day</td>
<td>351 days</td>
<td>No (promised with publication)</td>
</tr>
<tr>
<td>Toolformer</td>
<td>Yes, NeuRIPS</td>
<td>Yes</td>
<td>$770 \times$ 1.5 times a day</td>
<td>455 days</td>
<td>Yes</td>
</tr>
<tr>
<td>Leak, Cheat, Repeat</td>
<td>Yes, EACL</td>
<td>Yes</td>
<td>$12 \times$ Once in 8 days</td>
<td>93 days</td>
<td>Yes</td>
</tr>
<tr>
<td>Mistral 7B</td>
<td>No</td>
<td>No</td>
<td>$175 \times$ Almost daily</td>
<td>212 days</td>
<td>Inference only</td>
</tr>
<tr>
<td>LangBridge</td>
<td>No</td>
<td>No</td>
<td>$3 \times$ Once in 33 days</td>
<td>111 days (close to ACL deadline)</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Reading Papers

- Do read papers as primary sources
- Be critical, think about strengths and weaknesses of the papers
- Check for context (other benchmark numbers, publishing venue, reviews, citations, who wrote it)

https://ufal.cz/courses/npfl140