

Text-generation with rhyme and rhythm

Ans Farooq 2390370f

December 17, 2021

1 Proposal

1.1 Motivation

Poetry presents an additional challenge on top of text-generation in Natural Language Processing as it relies on the rhyme and rhythm of language. Factoring in these additional constraints of poem structure, rhyme, rhythm, number of lines and the lengths of lines presents an interesting challenge to new deep learning-based models.

1.2 Aims

This text-generation project will examine the use of transformer-based deep learning methods and the addition of length, rhyme and rhythm, given a certain topic for the poem to relate to. The end goal will be to produce cohesive and short limericks which should make some sense, be poetic, coherent and hopefully provide a bit of humour. The quality of these limericks will be measured by users.

2 Progress

- Current prototype combines a causal language model and a masked language model, GPT-2 and RoBERTa, to generate the starts of sentences and fill in the rest until the end rhyming word.
- A Wikipedia summary of the user's chosen topic is obtained and used as a prompt for GPT-2 to generate the lines in the limerick.
- The first and third lines of the limericks are generated repeatedly until a line with a decent end word is produced.
- A filter list is used to filter out lines generated by GPT-2 that have ending words that don't rhyme.
- NLTK cmudict (Natural Language Toolkit: Carnegie Mellon Pronouncing Dictionary) is used to find the end rhyming words.
- A list of the 5000 most common words is used to filter out uncommon/rare/obscure rhyming words that could be generated by NLTK.
- Added a full stop to the end of the sentences when filling in blank words with RoBERTa, improving output.
- The demo is hosted on HuggingFace spaces using Gradio.

3 Problems and risks

3.1 Problems

The following issues were encountered in the project so far.

- Finding out how to improve the rhyme-finding took a lot of time as most Python libraries produced very obscure, inconsistent results most of the time. Current state took a lot of experimenting.
- A lot of time was spent experimenting to get the program to generate coherent sentences relevant to the given topic. Current state is a stark improvement but still not good enough.
- On the odd occasion, the output generated is very random and seems to defy the programmed rules for limerick generation. Struggling to determine cause.

3.2 Risks

- Output from GPT-2 not good enough to produce coherent, relevant-to-topic sentences, never mind the sentences sounding poetic. **Mitigation:** Try using a different language model (GPT-3?), and experiment with the Wikipedia summary prompts.
- Limerick generation takes a very long time, and currently using GPT-2 instead of something more powerful. **Mitigation:** Explore how HuggingFace spaces and Gradio might help improve performance and allow for the use of GPT-3. Apart from this, no clear mitigation available.
- Unclear how to evaluate the success of the project. **Mitigation:** Investigate user surveys and ratings of the limericks generated? Develop a methodology of ranking the performance/quality of the limerick generation. Discuss with supervisor.

4 Plan

4.1 Semester 2

- Week 1-2: Use HuggingFace spaces. Focus on going from formal to poetic writing by experimenting with and implementing best approaches for syllables and rhymes. Investigate stressed/unstressed syllables, number of syllables in words and pronunciation. **Deliverable:** A (hopefully) improved system of limerick generation that is more poetic with cohesive English.
- Week 3-5: Develop limerick ranking method that can be used to generate multiple options, rank them, and pick the best one. Also continue improving limerick generation. **Deliverable:** Limerick ranking system that can be used for evaluation.
- Week 6: Simultaneously begin dissertation write-up and research how to best evaluate performance of final system. Also continue improving limerick generation, if possible. **Deliverable:** Start of dissertation and detailed evaluation plan, with participant numbers, information sheet and analysis plan.
- Week 7-9: Simultaneously continue dissertation write-up and complete final implementation and final improvements to limerick generation. **Deliverable:** Polished limerick generation system ready, passing basic tests, ready for evaluation stage.

- Week 9: Evaluation experiments run. **Deliverable:** Quantitative and qualitative measures of the quality of limericks generated and the performance of the limerick generation system.
- Week 8-10: Complete dissertation write up. **Deliverable:** First draft submitted to supervisor two weeks before final deadline.

5 Ethics

This project will involve tests with human users. These will be user studies using standard hardware, and require no personally identifiable information to be captured. I have verified that the experiments I plan to do comply with the Ethics Checklist.