University *of Ljubljana*
Faculty *of Computer and Information Science*

Conversations with Characters in Stories for Literacy

Žan Kogovšek, Žiga Drab, Vid Cesar

Abstract

In addressing the challenge of declining literacy interest among youth, we examine the use of character-driven chatbots for educational enhancement, aiming to make reading more engaging. We explore related works and existing NLM-based services, highlighting their strengths in user engagement and identifying inaccuracies in content representation. Our future work involves developing a framework for generating and evaluating such chatbots, with a focus on accuracy, engagement, and educational value.

Keywords

Literacy Improvement, Dialogue Agents, Harry Potter Dialogue Dataset, Pre-trained Language Models

Advisors: Slavko Žitnik

Introduction

The diminishing interest in reading among young individuals and their struggle with high-level literacy skills present a significant challenge to their academic and professional success [1]. Given literacy's crucial role, innovative methods are needed to motivate the young to engage with literature.

One possible solution is through the innovative use of chatbots. The research conducted by Brandtzaeg et al. [2] highlights that one of the primary reasons for the adoption of chatbots lies in their entertainment value. This finding suggests that chatbots serve not just as tools for information retrieval or customer service, but also as platforms for entertainment.

Lately, the evolution of open-domain conversation models, Adiwardana et al. [3], has achieved remarkable progress, uncovering the possibility of human-like sensibleness in machine conversations. Models can overcome the limitations of simple exchanges by incorporating traits commonly found in real-life interactions, such as emotion [4] and empathy [5]. This approach, often referred to as style-controlling conversation models, significantly enhances the quality of engagement, making the conversation feel more genuine and relatable [6]. Zhang et al. [7] introduce a model capable of producing responses with consistent personalities by utilizing personal descriptions, a step forward in making conversation models more personalized and engaging.

In this paper, we explore the potential of enhancing children's literacy through entertaining chatbots with personalities, modeled after characters from literature. Our objective is to develop a streamlined pipeline capable of generating chatbots across various datasets, thereby enabling children to interact with their favorite figures and make reading more appealing.

Related Work

Enhancing children's literacy through technology and innovative approaches has been already explored by Nielen et al. [8]. They introduced a pedagogical agent, an animated mouse guiding fifth graders in their reading. This agent boosted reading motivation and vocabulary learning by engaging them with summaries and reflective questions.

Moving on to personalized dialogue agents, Chen et al. [9] introduced the Harry Potter Dialogue (HPD) dataset, which contains not only dialogues but also background information, including scenes, speakers and their relationship with Harry. They evaluated LLMs, like Alpaca, ChatGPT and GPT-3 on the HPD dataset through fine-tuning and in-context learning. Results indicate significant potential for improvements, yet the dataset proves useful in steering models towards more accurately reflecting Harry Potter's character.

Han et al. [10] introduce a strategy, Pseudo Dialog Prompting (PDP), that utilizes statements from various characters alongside a pre-trained language model to generate conversations. Among the pre-trained models selected for this purpose are GPT-2-XL, GPT-J, and GPT-Neo. The performance of these models was tested both through automatic evaluation using MaUdE and via human assessments. The findings from these experiments demonstrate that the PDP strategy successfully crafts responses that accurately mirror the style of the specified characters.

Figure 1 presents a generic framework demonstrated by Hefny et al. [11] for creating character-based chatbots structured around five modules. These include the conversational user interface, entities, intents, webhook, and database. The conversational user interface serves as the initial point of interaction, allowing users to input prompts and select entities. Following this, the system loads the chosen entity and attempts to understand the user's intent. This intent is then processed through a webhook module, which facilitates access to the database. From here, responses are randomly selected and presented to the user

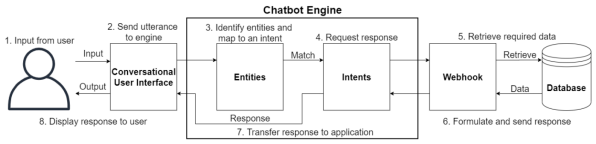

Figure 1. Proposed generic framework for building character-based chatbots [11].

Bot Persona Services

We tested Character.ai, a service offering interactive dialog agents based on Neural Language Models (NLM) for entertainment. We began a conversation with Greek Mythology hero Achilles (Version 1 Achilles). The responses generated were not just the responses of a character, but also included his actions - "nodding slowly" for example, and even interpretation of said actions - "expression full of sadness and regret". The model was reluctant to give out information about himself, which he deemed important, and was trying to find out what we knew about him. When faced with direct questions about something, it seemed as if it was trying to avoid answering it - maybe because of the character's nature. Testing the model further, we noticed incorrectness in story of his death. Despite repeated inquiries, the model persisted with the incorrect narrative, highlighting concerns over its handling of basic factual accuracy.

We also tried another model of the same character. While avoiding the previous error, it inaccurately identified Achilles' parents, mistaking his father for a god and his mother for a human. The model again struggled with factual accuracy. Despite a disclaimer on Character.ai indicating that responses are fictional, the inaccuracies observed in the models' responses raise concerns about factual correctness.

For improving literacy, some dialog agents can be used as they use rich vocabulary and there is a lot of extra text in addition to the answer itself. Yet, some agents provide extremely short answers. Example of such dialog agent is Miles Morales (Version 1 Miles Morales), where even information about character actions and context is very sparse, indicating not all agents are effective for literacy improvement. Model descriptions, using terms like "cold" or "rude," can also guide the selection of agents supportive to literacy goals.

Initial Idea

Adapting the generic framework for constructing character-based chatbots, we simplify it by removing the database and webhook modules. These are replaced with an existing model capable of generating responses to prompts in real time. Our revised framework retains a user interface that enables users to select their preferred entity. Additionally, an entity selector is employed to choose the suitable character model from the character model pool. The architecture of our updated framework is illustrated in Figure 2.

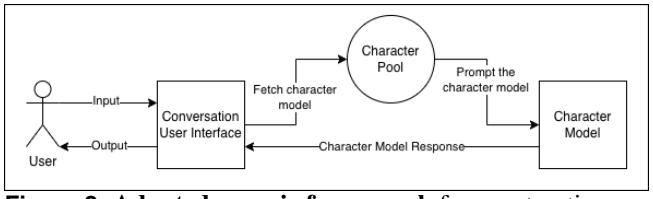

Figure 2. Adapted generic framework for constructing character-based chatbots with 3 modules.

Our framework requires a diverse collection of character models, comprising the character model pool. The process for assembling this pool is detailed in Figure 3. Initially, the dataset pool will include the Harry Potter dataset mentioned earlier, with plans to expand it with additional datasets. The Pre-Trained Model pool will feature GPT-2 models, as discussed in the related works and will also include the Llama pre-trained model if possible. These datasets and pre-trained models will be merged through supervised fine-tuning to create character-specific models, which are then gathered in the character model pool. Given the complexity of these pre-trained models, which consist of billions of parameters, the creation and management of multiple character chatbots will necessitate the use of high-performance computing resources.

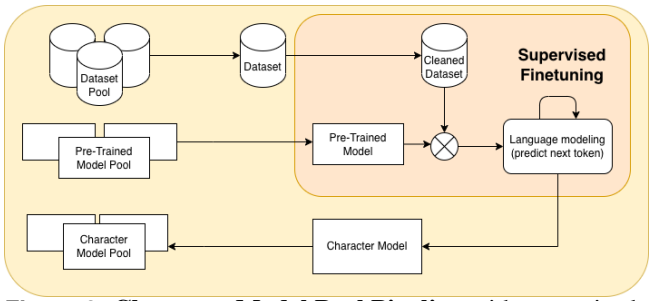

Figure 3. Character Model Pool Pipeline with supervised finetuning.

References

[1] Jane Murray. Literacy is inadequate: young children need literacies, 2021.

[2] Petter Bae Brandtzaeg and Asbjørn Følstad. Why people use chatbots. In *Internet Science: 4th International Conference, INSCI 2017, Thessaloniki, Greece, November*

22-24, 2017, Proceedings 4, pages 377–392. Springer, 2017.

[3] Daniel Adiwardana, Minh-Thang Luong, David R So, Jamie Hall, Noah Fiedel, Romal Thoppilan, Zi Yang, Apoorv Kulshreshtha, Gaurav Nemade, Yifeng Lu, et al. Towards a human-like open-domain chatbot. *arXiv preprint arXiv:2001.09977*, 2020.

[4] Hao Zhou, Minlie Huang, Tianyang Zhang, Xiaoyan Zhu, and Bing Liu. Emotional chatting machine: Emotional conversation generation with internal and external memory. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

[5] Hannah Rashkin, Eric Michael Smith, Margaret Li, and Y-Lan Boureau. Towards empathetic open-domain conversation models: A new benchmark and dataset. *arXiv preprint arXiv:1811.00207*, 2018.

[6] Eric Michael Smith, Diana Gonzalez-Rico, Emily Dinan, and Y-Lan Boureau. Controlling style in generated dialogue. *arXiv preprint arXiv:2009.10855*, 2020.

[7] Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. Personalizing dialogue agents: I have a dog, do you have pets too? *arXiv preprint arXiv:1801.07243*, 2018.

[8] Thijs MJ Nielen, Glenn G Smith, Maria T Sikkema-de Jong, Jack Drobisz, Bill van Horne, and Adriana G Bus. Digital guidance for susceptible readers: Effects on fifth graders' reading motivation and incidental vocabulary learning. *Journal of Educational Computing Research*, 56(1):48–73, 2018.

[9] Nuo Chen, Yan Wang, Haiyun Jiang, Deng Cai, Yuhan Li, Longyue Wang, Jia Li, et al. Large language models meet harry potter: A dataset for aligning dialogue agents with characters. In *The 2023 Conference on Empirical Methods in Natural Language Processing*, 2023.

[10] Seungju Han, Beomsu Kim, Jin Yong Yoo, Seokjun Seo, Sangbum Kim, Enkhbayar Erdenee, and Buru Chang. Meet your favorite character: Open-domain chatbot mimicking fictional characters with only a few utterances. *arXiv preprint arXiv:2204.10825*, 2022.

[11] Walid El Hefny, Alia El Bolock, Cornelia Herbert, and Slim Abdennadher. Towards a generic framework for character-based chatbots. In *Highlights in Practical Applications of Agents, Multi-Agent Systems, and Trustworthiness. The PAAMS Collection: International Workshops of PAAMS 2020, L'Aquila, Italy, October 7–9, 2020, Proceedings 18*, pages 95–107. Springer, 2020.