

SciFiGPT: Finetuning GPT-2 for Short Science Fiction Generation

Yu Xia

University of Michigan

xiayuu@umich.edu

1 Introduction

The field of natural language processing (NLP) has seen significant progress in recent years, thanks to advanced language models such as GPT-2. These models have the ability to generate coherent and fluent text, which presents exciting possibilities for creative writing and content creation. Among the many applications of NLP, generating science fiction stories is a particularly challenging and interesting task due to the genre's emphasis on imagination and creativity.

In this project, we aim to finetune the GPT-2 model using a large corpus of science fiction text to produce high-quality, imaginative, and coherent stories. Our objective is to create science fiction stories that are indistinguishable from those written by human authors. This project differs from previous approaches by focusing on the science fiction genre and using a large corpus of text to improve the quality of the generated stories.

Solving this problem would have important implications for the NLP community, as well as practical applications for creative writing, content creation, and technology industries. The ability to generate high-quality science fiction stories could benefit authors, editors, and publishers by improving the speed and efficiency of content creation. Additionally, it could have applications in the entertainment industry, where the demand for unique and compelling stories is high.

Our proposed method for solving this problem involves fine-tuning the GPT-2 model (SciFiGPT) using a large corpus of science fiction text and evaluating the quality and coherence of the generated stories using various metrics. The results of this project will be of interest to NLP researchers, creative writers, and technology companies. By demonstrating the potential of advanced language models to generate high-quality science fiction stories, we hope to inspire further research and devel-

opment in this area.

In summary, this project aims to demonstrate the ability of the GPT-2 model to generate high-quality science fiction stories through fine-tuning and evaluation. The contributions of this project include insights into the capabilities of advanced language models for creative writing and content creation, as well as a deeper understanding of the challenges and opportunities presented by the science fiction genre.

2 Data

For this project, we used the science fiction stories text corpus available on Kaggle. This dataset was created by Jannes Klaas and contains a collection of over 100 science fiction stories from the 1950s to the 1960s. The stories were obtained from Project Gutenberg and have been preprocessed to remove headers and footers.

Here are some rough statistics about the dataset before preprocessing:

- Number of words: 1,072,830
- Number of sentences: 2,212,421
- Number of paragraphs: 2,212,42

The dataset is provided as a single text file. To prepare the data, we split the data into training and validation subsets for fine-tuning and evaluating our model. We used the `TextDataset` class from the `transformers` library to load the dataset, tokenize it using the GPT-2 tokenizer, and split it into training and validation sets. We set the `block_size` parameter to 128 to ensure that the data is split into sequences of manageable length for training.

Here are some rough statistics about the dataset after preprocessing:

- Training dataset size: 243,564
- Validation dataset size: 27,063

Example story from the dataset:

Title: The Variable Man Author: Philip K. Dick

He fixed things—clocks, refrigerators, vidsenders and destinies. But he had no business in the future, where the calculators could not handle him. He was Earth’s only hope—and its sure failure!

Security Commissioner Reinhart rapidly climbed the front steps and entered the Council building. Council guards stepped quickly aside and he entered the familiar place of great whirring machines. His thin face rapt, eyes alight with emotion, Reinhart gazed intently up at the central SRB computer, studying its reading.

3 Related Work

The field of text generation using deep learning models has seen significant work, with various approaches such as recurrent neural networks (RNNs) and transformers. OpenAI’s GPT-3 (Brown et al., 2020) is a large-scale transformer-based language model that was trained on a massive corpus of text data. GANs (Nie et al.; Yang et al., 2020) have been used to generate realistic and diverse text samples. For this project, we will use OpenAI’s open-sourced GPT-2 (Radford et al., 2019) as the base model for our text generation task.

Several studies have explored text generation in different domains. For instance, (Keskar et al., 2019) introduced a method for controlling the style and content of generated text by conditioning the model on control codes. (Dathathri et al., 2019) proposed a plug-and-play method for fine-tuning GPT-2 on specific tasks without the need for task-specific training data.

Our approach differs from previous work in that we will fine-tune GPT-2 on a large corpus of science fiction text to generate imaginative and coherent stories. By focusing on a specific genre and providing the model with a vast amount of relevant training data, we hope to improve the quality and coherence of the generated stories. We believe that our approach has the potential to produce science fiction stories that are indistinguishable from those written by human authors.

We also note that while there are several studies on text generation, there is limited work on generating science fiction stories. Our project aims to

address this gap by generating stories in a specific genre that requires a high level of creativity and imagination. We hope that our work will contribute to the understanding of the potential of language models for creative writing tasks.

4 Method

4.1 Data Preprocessing

To prepare the data for fine-tuning the GPT-2 model, we first tokenize and encode the input text using the TextDataset class from the Hugging Face Transformers library (Wolf et al., 2020). We provide the TextDataset class with the tokenizer object and the file path to the science fiction text data. We also set the block size to 128, which determines the maximum length of the input sequences used for training the model.

We then split the data into training and validation subsets using the torch.utils.data.random_split function, with 90% of the data used for training and the remaining 10% used for validation. We also print the number of samples in each subset to ensure that the split is correct.

4.2 Model Finetuning

To fine-tune the pre-trained GPT-2 language model (Radford et al., 2019) on the science fiction text corpus, we use the Trainer class from the Transformers library. We provide the Trainer class with the fine-tuned GPT-2 model, the training and validation datasets, the data collator for language modeling, and the training arguments.

The data collator is defined using the DataCollatorForLanguageModeling class from the Transformers library, which collates the input sequences and the corresponding labels for each batch. We set the mlm parameter to False because we are not using masked language modeling for our task.

The training arguments are defined using the TrainingArguments class from the Transformers library, which specifies the hyperparameters for training the model. We set the output directory to ./results and the logging directory to ./logs. We also set the evaluation strategy to 'epoch', which evaluates the model after each epoch, and the number of training epochs to 3. We set the batch size to 50 for both training and evaluation, the warmup steps to 500, and the weight decay to 0.01. Finally, we set the logging steps to 1000 and the save total limit to 1.

After setting up the Trainer, we call its train method to start the training process. During training, the model generates sequences of text given the input sequences and the corresponding labels, and the optimizer updates the model’s parameters based on the loss between the generated sequences and the labels. The Trainer logs the training and validation metrics, including the loss, perplexity, and learning rate, to the logging directory. At the end of training, the Trainer saves the fine-tuned GPT-2 model to the output directory. And we refer the finetuned model as SciFiGPT for the rest of the report.

4.3 Evaluation and Results

To evaluate the performance of our model, we measured the perplexity score and conducted a human evaluation of generated text.

4.3.1 Perplexity Results

We measured the perplexity score of the model before and after fine-tuning. The results are presented in Table 1.

Table 1: Perplexity Scores

Model	Perplexity
GPT2	62.86
SciFiGPT	42.37

We can see that the perplexity score improved significantly after fine-tuning the model.

4.3.2 Human Evaluation Results

To evaluate the quality of the generated text, we used three three-sentence prompts and asked 5 human evaluators to rate the quality of the generated text on a scale of 1 to 5, where 1 is poor quality or not relevant to science fiction and 5 is excellent quality and relevant to science fiction. We used the model before and after fine-tuning to generate text for each prompt.

The three-sentence prompts are as follows:

1. The year is 2050 and humanity has just discovered a new planet capable of sustaining life. A team of scientists and explorers are sent to investigate. What they find will change the course of human history forever.
2. In a world where technology has advanced to the point of being able to upload human consciousness into a virtual reality, one man

must navigate this new world and the dangers that come with it. But when he discovers a dark secret about the virtual world, he must decide whether to stay in blissful ignorance or risk everything to expose the truth.

3. After a catastrophic event leaves Earth uninhabitable, the last remaining humans are forced to live on a space station orbiting the planet. But when a group of rebels discover a way to potentially save Earth, they must decide whether to risk their lives for a chance at redemption or continue living in their artificial home.

The generated texts of GPT2 and SciFiGPT are provided in Appendix due to space limitations. The results of the human evaluation of generated science fiction stories are presented in Table 2.

Table 2: Average Scores

Prompt	GPT2	SciFiGPT
#1	1.8	2.8
#2	1.2	2.8
#3	1.0	2.2

From the human evaluation results, we observe that SciFiGPT outperforms GPT2 in generating science fiction stories. However, the absolute quality of SciFiGPT is low as judged by human.

5 Discussion

The results of perplexity and the human evaluation of the generated science fiction stories show that SciFiGPT performs better than the baseline GPT2 in generating science fiction stories. However, the absolute quality of the generated stories is low as judged by human evaluators. This suggests that there might be certain issues with the training data or the training process.

One possible reason for the low quality of the generated stories could be the insufficient amount of training data. The training data might not be representative enough of the entire space of possible science fiction stories, leading to the model not being able to capture the nuances and intricacies of the genre. Another possible reason could be the low quality of the training data, which might be affecting the overall performance of the model. The training data might have inconsistencies or biases, leading to the model not being able to generate high-quality stories. As a result, the difference in

performance of the model before and after finetuning is not as significant as we might have expected.

Overall, the performance of our approach is satisfactory for an end-user of our model. The results suggest that more work needs to be done to improve the quality of the training data and to explore more advanced techniques or use a larger model to further improve the performance.

6 Conclusion

In this project, we aimed to demonstrate the ability of the GPT-2 model to generate high-quality science fiction stories through fine-tuning and evaluation. We developed an approach to generate science fiction stories by finetuning the GPT-2 model using a large corpus of science fiction text. Our experiments showed that the finetuned model (SciFiGPT) outperformed the baseline GPT-2 in terms of perplexity and human evaluation metrics. However, we found that the absolute quality of the generated stories was low, indicating that there is scope for improvement in the training data or the training process. One possible reason for this is the insufficient amount of training data, which might not be representative of the entire space of possible science fiction stories. Another reason could be the low quality of the training data. Future work could involve increasing the amount of training data and exploring more advanced techniques to improve the performance of the model. Additionally, using a larger model, such as GPT-3 or GPT-Neo, could potentially lead to better results. Another direction could be to experiment with different preprocessing techniques or evaluate the model's performance on specific sub-genres of science fiction.

7 Other Things We Tried

During the course of this project, we attempted to train a T5-small model on the dataset. However, we encountered challenges during the preprocessing phase, as the dataset is in the format of single-line text file consisting of a large corpus of science fiction stories. It could not be easily converted into the input-to-target format required by the sequence-to-sequence (encoder-decoder) architecture used by T5-small. After spending some amount of time trying to preprocess the data for use with the T5-small model, we eventually had to abandon this approach due to the time constraints.

8 What You Would Have Done Differently or Next

if we were to do this project again, we would have chosen a different downstream task that could be easily addressed by both GPT-2 and T5-small models. For example, we could have chosen the task of text summarization, which involves generating a shorter version of a given text while retaining its most important information.

A SciFiGPT Generated Stories

A.1 Prompt 1

The year is 2050 and humanity has just discovered a new planet capable of sustaining life. A team of scientists and explorers are sent to investigate. What they find will change the course of human history forever. "A giant man-sized creature with tentacles, arms, legs, eyes – we call him Man." The words had struck me as such a wild guess that I didn't have time for them. But it was true. They were there now and not far away. And even at this distance... At least in their own minds, which were still reeling from the shock of my first vision, men like us could see things clearly enough. We found ourselves talking together: "It's been six years since you've told us about what happened before your death. How long?" "Four. Then here's where we're going." It wasn't tiring to be alone again; most of us lay asleep all night without sleep. All right, maybe if we hadn't thought of that, but who else would want to? There'd been no one around after our last visit? He looked up into the darkness outside and saw nothing. I felt so tired out he walked back toward the cave. No, his body shivered under his feet. His mind went crazy. He couldn't feel anything inside himself or touch any part of him. Nothing moved except the sky, the wind, and the soft blue glow of the forest. Suddenly he turned suddenly to face the trees. Something seemed different.

A.2 Prompt 2

In a world where technology has advanced to the point of being able to upload human consciousness into a virtual reality, one man must navigate this new world and the dangers that come with it. But when he discovers a dark secret about the virtual world, he must decide whether to stay in blissful ignorance or risk everything to expose the truth. # pp., List Price...# Discounts lor only I S-F is so

far beyond what we've seen on TV – but at least you're getting an idea for how imaginative science fiction can be! For just four dollars per issue send (# issues) by letter (check your subscription label), pay nothing; all cash payments will go directly to supporting The Foundation! ; plus several cents shipping charges.) IF Short Story L arthur was fascinated as well as disturbed over some vague hints which had been offered him concerning his first appearance on Earth after returning from the spacelings' colony planetoid. He found them too obvious, yet more puzzling than anything else. As they began to investigate their own strange surroundings, suddenly he saw there was something almost impossible about them. At once he realized why these creatures were no longer intelligent enough to communicate with others: if any part existed within itself, its purpose was utterly unknown. There was much room in space for speculation here, however; most likely such questions would have to be answered before long.

A.3 Prompt 3

After a catastrophic event leaves Earth uninhabitable, the last remaining humans are forced to live on a space station orbiting the planet. But when a group of rebels discover a way to potentially save Earth, they must decide whether to risk their lives for a chance at redemption or continue living in their artificial home. The book is one of those rare gems that have never been published before – and it's an important addition to any science fiction library! (Publ., #) by Robert Silverberg Theodore Sturgeon Isaac Asimov Harlan Ellison Damon Knight Frank Herbert Locus Press Edmond Hamilton Fritz Leiber Jack Vance Lester del Rey Frederik Pohl James Blish Roger Zelazny Larry Niven A S E J H OW was born into slavery during World War II. He came from a family of wealthy farmers who had bought up all the land under the New York Stock Exchange as slaves; but his parents were able only to sell them out because he could not find enough money to buy back what he needed. When he reached college, he spent most of his time working with the New York Stock Exchange. In school, he learned how to trade paper currency so that people would think about him. His interest in trading became more than just a hobby; he found himself involved with other men and women whose ideas might help others gain knowledge. After graduation, he joined

the American Association of Certified Professional Engineers, which made training possible for both professional engineers and mechanics.

B GPT Generated Stories

B.1 Prompt 1

The year is 2050 and humanity has just discovered a new planet capable of sustaining life. A team of scientists and explorers are sent to investigate. What they find will change the course of human history forever. This movie gives you more than 10 hours (20 minutes per episode) that tells your story, explore what happens in space but not where it's going: through intergalactic journey by astronaut Dr. Halsey; encounters with other characters from Halo 3 - all this while remaining true-to form for fans who love gameplay rather much! And be sure there's also an option called "Secret Agent" which allows players get free access into their favorite areas without any restrictions or limitations on how many missions can appear at once within one hour so long as those special objectives don't go unmet if no mission exists before then...

B.2 Prompt 2

In a world where technology has advanced to the point of being able to upload human consciousness into a virtual reality, one man must navigate this new world and the dangers that come with it. But when he discovers a dark secret about the virtual world, he must decide whether to stay in blissful ignorance or risk everything to expose the truth. I love working at Google because I have always enjoyed creating original content for my team-mates so much but can you tell us more? In your opinion what are some things we've learned from our experience on Kickstarter? A lot - how do people feel seeing their work as something they should be proud of then getting funded by other projects rather than just donating money back up front (for example) is important too. Do others find these types of funding an interesting way of helping them make sure great art comes out before crowdfunding all together again! Tell me any story like yours, please share ;-)

B.3 Prompt 1

After a catastrophic event leaves Earth uninhabitable, the last remaining humans are forced to live on a space station orbiting the planet. But when a group of rebels discover a way to poten-

tially save Earth, they must decide whether to risk their lives for a chance at redemption or continue living in their artificial home. A New Hope is an original story about human survival and hope that takes place over two decades after humanity’s destruction by war against aliens who have been hiding from mankind since before our own time (and we’re pretty sure this isn’t going anywhere). It follows four astronauts – Matt Damon, Mark Watney; Robert Downey Jr., Anthony Rapp, John Chopper & James Franco – as one man trying desperately not only survive but also become part-human again with what may well be his first mission back into outer space!

References

- Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. [Language models are few-shot learners](#). In *Advances in Neural Information Processing Systems*, volume 33, pages 1877–1901. Curran Associates, Inc.
- Sumanth Dathathri, Andrea Madotto, Janice Lan, Jane Hung, Eric Frank, Piero Molino, Jason Yosinski, and Rosanne Liu. 2019. Plug and play language models: A simple approach to controlled text generation. *arXiv preprint arXiv:1912.02164*.
- Nitish Shirish Keskar, Bryan McCann, Lav R Varshney, Caiming Xiong, and Richard Socher. 2019. Ctrl: A conditional transformer language model for controllable generation. *arXiv preprint arXiv:1909.05858*.
- Weili Nie, Nina Narodytska, and Ankit Patel. Relgan: Relational generative adversarial networks for text generation. In *International conference on learning representations*.
- Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.
- Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander Rush. 2020. [Transformers: State-of-the-art natural language processing](#). In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45, Online. Association for Computational Linguistics.
- Yang Yang, Xiaodong Dan, Xuesong Qiu, and Zhipeng Gao. 2020. [Fggan: Feature-guiding generative adversarial networks for text generation](#). *IEEE Access*, 8:105217–105225.