



HeartLink: A Love Letter Generator

Final Report

Word count: 1956

Penalty: 0%

Hexuan Zhang | 1010737909

Hanxiao Wei | 1010648792



Permissions

All Members of the *HeartLink* group grant limited permissions for the ECE1786 instruction team to post the following items on a public course website:

- Permission to post video: Yes
- Permission to post final report: Yes
- Permission to post source code: No

Hexuan Zhang

Hanxiao Wei



Introduction

Our project "*HeartLink*" is a love letter generator. It's a tool designed for those moments when feelings are strong, but words seem just out of reach. By simply providing a **Name**, a **Relationship**, and some **Descriptions of Special Moments**, *HeartLink* will craft a love letter for you, echoing the timeless charm of classic love tales. Our project is interesting, fun to work on, and fun to use.

Illustration & Figure

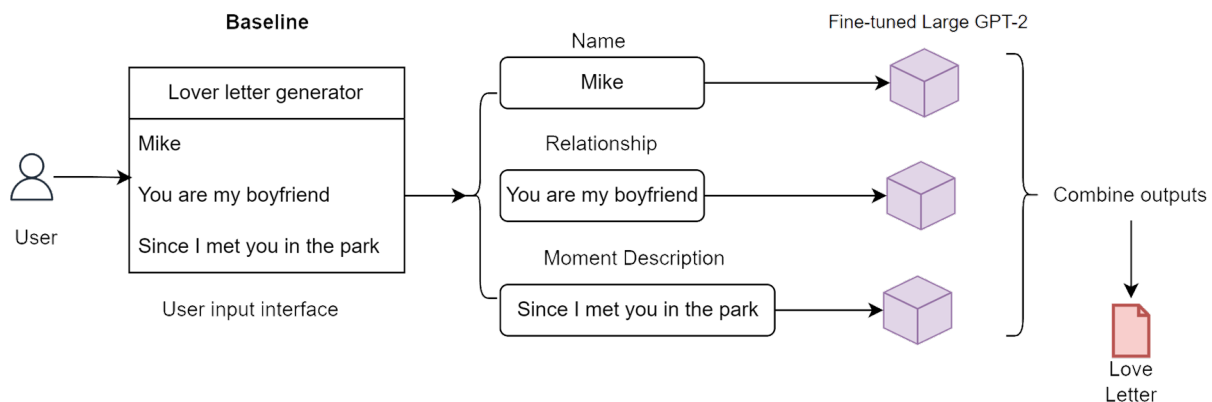


Figure 1. Baseline System Architecture

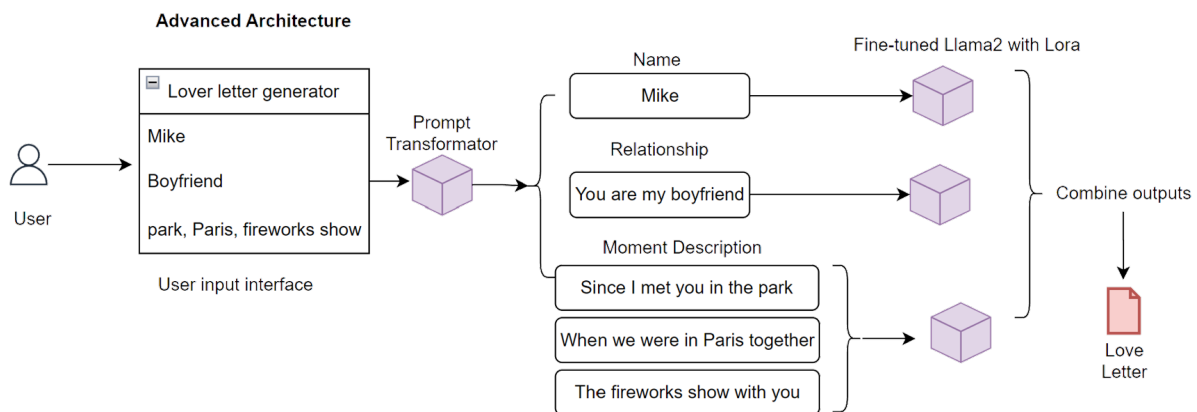


Figure 2. Main (Advanced) System Architecture



Background & Related Work

The advanced NLP models have opened up new possibilities in creative text generation. However, there appears to be a limited number of studies focused on the generation of love-letter themed content.

- Intelligent Love Letter Generator Based on GPT-2 Model [1]

This paper shows the implementation of a love letter generator using the GPT-2 model, and the training of a LSTM sentiment classifier. However, the focus of this paper is more about the LSTM part. This is the only document specifically related to generating love letters or love words that we found publicly available online, while it is quite different from what we are trying to do in our project, which could emphasize the uniqueness of our topic selection.

- Automatic Story Generation: Case Study of English Children's Story Generation Using GPT-2 [2]

This paper explores the potential of GPT-2 in creating English children's stories. It evaluates the model's capability in crafting coherent narratives.

Data Processing

Our Raw Dataset contains 2000+ sentences collected from various license-free sources:

- We downloaded 87 love letters and segmented them into individual sentences.
- We copy-pasted hundreds of love messages from various websites.
- We used GPT4 to generate 35 sentences.

We have performed various meticulous to improve the dataset quality:

- We removed irrelevant content, symbols, and special punctuations.
- We removed specific names.
- We filtered out sentences with more than 30 words and rewrote them into shorter sentences.
- We removed extremely short, or non-complete sentences.
- We removed inappropriate, offensive, or sensitive content.
- We manually checked the quality of each sentence.

We didn't need to split our dataset. All data is training data. Our final processed dataset contains 1487 lines of sentences, in the range of 3 to 30 words per sentence. The word count distribution is shown below in Figure 3.

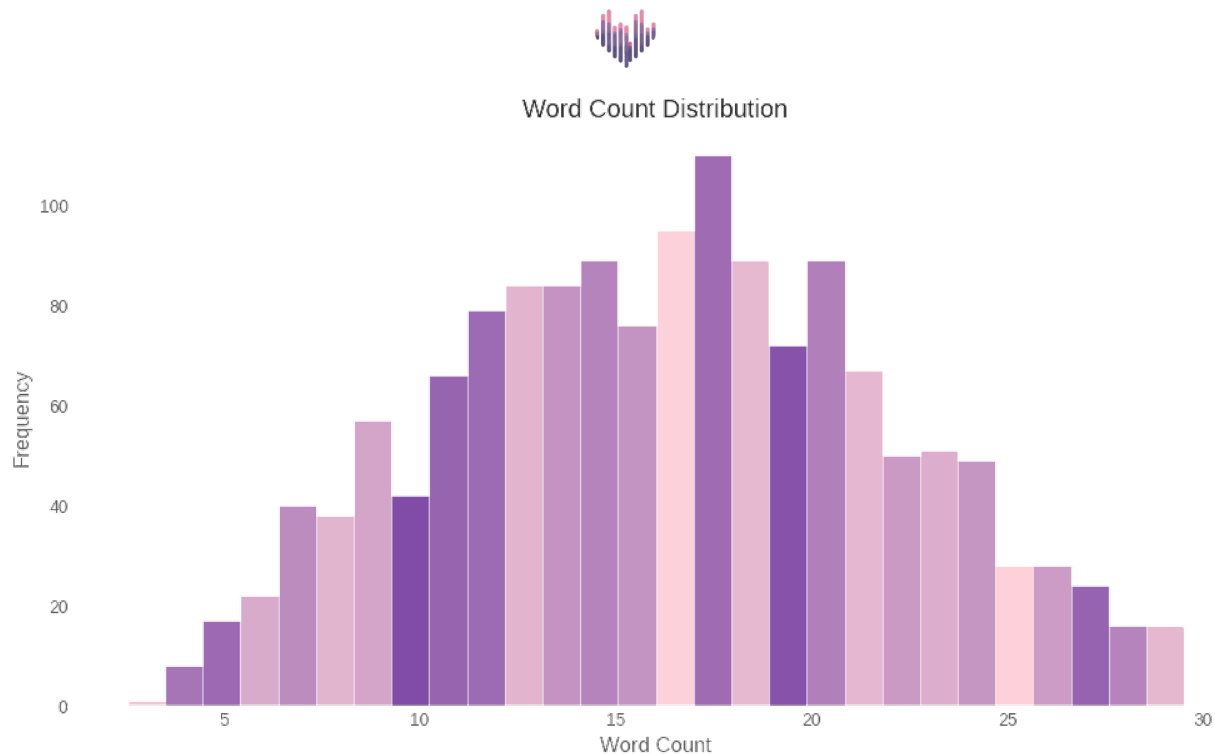


Figure 3. Word Count Distribution of Processed Dataset

One example of the cleaned training data:

Love is a deep friendship, understanding, and loyalty, accepting imperfections and forgiving weaknesses.

Main System

In our main system, we used the Llama2-7b model to produce sentences that are both longer and more cohesive. Due to the substantial computing resources needed to finetune the Llama2-7b model, we attempted over twenty times to access the A100 GPU with 80GB of memory on Colab Pro but were unsuccessful in securing the A100 even once. Drawing inspiration from a blog post [3], we explored a method to fine-tune the Llama2-7b model using the PEFT (Parameter-Efficient Fine-Tuning) library. This approach utilizes the QLoRA (Quantized Low Rank Adapters) technique, which allows us to fine-tune adapters that are placed on top of the frozen 4-bit model. The key advantage of using LoRA (Low Rank Adapters) is that it makes only a small portion of the model trainable. As a result, both the number of learned parameters and the size of the trained artifact are significantly reduced. This efficient method enables us to fine-tune the Llama2-7b model on a T4 GPU, which has only 16GB of memory, making it a more feasible and resource-efficient approach for model training.

The training curve of our main model is shown below in Figure 4.

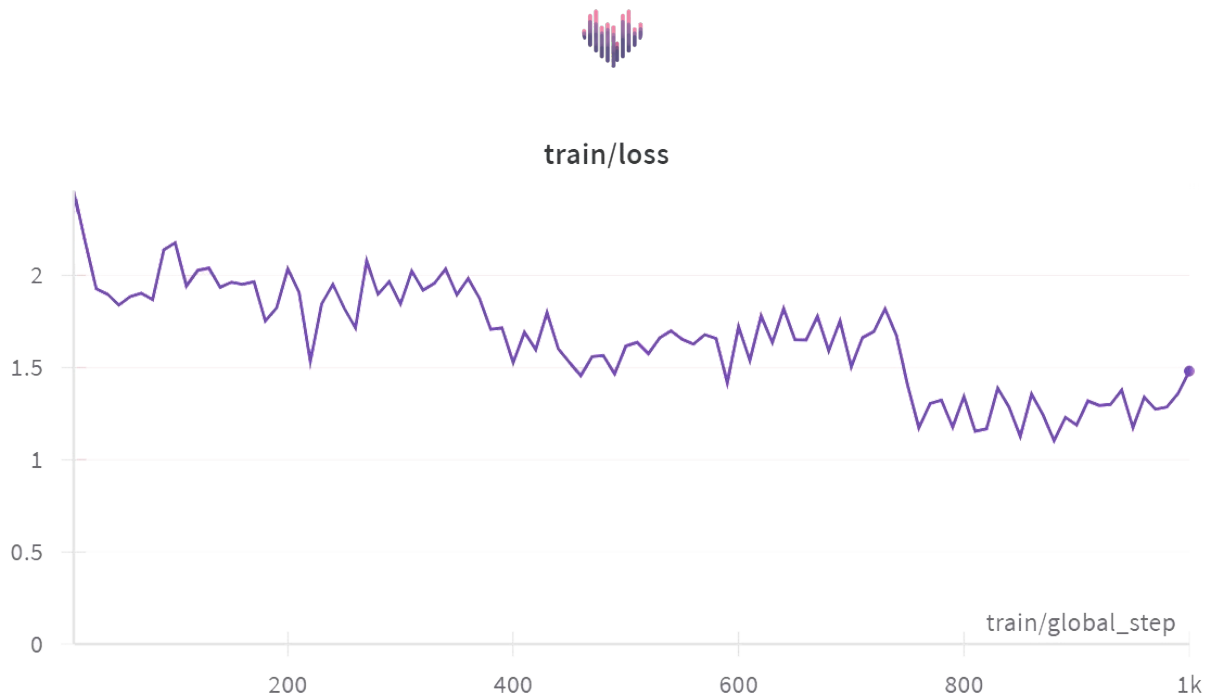


Figure 4. Main System Training Loss Curve

The *Transformator* Module

Sometimes, user Inputs may not be immediately suitable to start a sentence. So, we introduced a new Prompt *Transformator* module. This module is designed to convert user inputs into well-formulated prompts that align with the model's input requirements. It works as follows:

Initial Check: Determines if the input is a single word, a phrase, or a sentence.

Pronoun/Auxiliary Verb Check: For multi-word inputs, it checks for pronouns or auxiliary verbs to identify complete sentences. No modification is needed for complete sentences.

Check Gift Type: The function identifies if the input matches popular gifts, pets, car brands, luxury brands, or travel-related terms using a high-similarity threshold. If there's a match, the input is transformed with a predefined string format.

Check Word Type: The function utilizes NLTK to determine the word type (verb, adjective, adverb, etc.). If it's a noun, the function will additionally use dictionaries to see if that's a place. The input is then transformed with a specific string format based on its word type.

This module is equipped exclusively on the main model.

Baseline System

Our project has a baseline system that utilizes the same dataset for fine-tuning as the main system. However, the baseline system does not have the *Transformator* module, and employs GPT2-Large model instead of the Llama2-7b.



The training curve is shown below in Figure 5.



Figure 5. Baseline System Training Loss Curve

Qualitative Results

Figure 6 shows the user interface of our project. It is beautifully designed, which could be a qualitative highlight of our project.

HeartLink: A Love Letter Generator

HeartLink will craft a letter-style love message for you, echoing the timeless charm of classic love tales. Fill in the details and press "Generate a Love Letter" to create your personalized message.

Please input your lover's Name
Mike

Please input your relationship
boyfriend

Please input a few words or phrases about the moments you love, each separated by a comma
park, Paris, fireworks show

Generate My Love Letter

Your Love Letter

Mike, you are the love of my life, a man who has shown me what it means to love and be loved. I cherish you and your love. You are my boyfriend, and I love you. I don't know how you can be so cute and so handsome. Since I met you in the park, I've felt as though we were destined to be together. My heart and mind have been filled with thoughts of you. When we were in Paris together, I had the sense that we had been in love for many lifetimes, and that we were simply reminiscing about our past lives. I believe in soulmates, and you're mine. The fireworks show with you is always the best. Bye, darling.

Figure 6. Example Program User Interface



In Figure 6, the love letter generated by the main system is (Text in **purple** is the user input):

Mike, you are the love of my life, a man who has shown me what it means to love and be loved. I cherish you and your love. You are my boyfriend, and I love you. I don't know how you can be so cute and so handsome. Since I met you in the park, I've felt as though we were destined to be together. My heart and mind have been filled with thoughts of you. When we were in Paris together, I had the sense that we had been in love for many lifetimes, and that we were simply reminiscing about our past lives. I believe in soulmates, and you're mine. The fireworks show with you is always the best. Bye, darling.

For comparison, the love letter generated by the baseline system is in Appendix 1.

Both systems perform better than our expectations. A significant factor contributing to the model's success is the considerable amount of time and effort we invested in sourcing high-quality training data, meticulously processing it, and conducting thorough checks to ensure its accuracy.

However, our systems did not perform well with some types of inputs:

- Due to the absence of a variable to regulate gender-specific descriptors, both systems are prone to gender coherence issues, which may result in outputs with contradictory gender references, such as: "You are my girlfriend, ... you are the best husband".
- Non-standard names may be misinterpreted as objects by our main model, e.g. Llama-2 may consider the input name 'Grape' to be some fruit, rather than a person's name. In contrast, this issue is not observed in the baseline model.
- If the input is a polysemy word that contains multiple meanings with different word types (e.g., noun and verb), the *Transformer* Module will be confused and interpreted it by a certain logic, which may cause misunderstanding or grammar issues.
- The main model may end the love letters in an odd way with certain inputs.

Quantitative Results

The first example of the love letter generated by the main model is available in the Qualitative Results section above. We evaluated the result based on the following 5 aspects, with 2 points for each aspect and a full mark of 10 points:

- Emotions: How powerfully the letter expresses feelings of love.



- Sincerity: How well emotional authenticity is presented.
- Personalization: How the letter is tailored to the specific person.
- Cohesion: How well the letter is structured and whether its message is easily understood, including grammar performance.
- Impacts: How effective the love letter could be on the recipient.

Each of us and GPT-4 graded love letters separately, and then we averaged the scores in each aspect.

Table 1: Love Letter 1 Evaluation

	Baseline Model	Main Model
Emotional Expression	1.67/2	1.93/2
Sincerity	1.73/2	1.93/2
Personalization	1.67/2	1.90/2
Clarity and Cohesion	1.40/2	1.16/2
Impact	1.16/2	1.67/2

The final score of the baseline system output is: 7.63/10, while the score for the main system is: 8.59/10. We think the main model content is a little redundant (e.g., lots of I love you) and not very cohesive (especially the end “Bye, darling”), but it is still better than the baseline result, overall.

The second example of the love letter generated by the main system is available in Appendix 2. We evaluated the second example with the same method:

Table 2: Love Letter 2 Evaluation

	Baseline Model	Main Model
Emotional Expression	1.80/2	1.93/2
Sincerity	1.67/2	1.93/2
Personalization	1.70/2	1.93/2
Clarity and Cohesion	1.63/2	1.70/2
Impact	1.60/2	1.83/2



The Average Score of the baseline model output is: 8.40/10. The Average Score of the main model output is: 9.32/10. The main love letter scores highly in all categories. It successfully conveys deep emotions and appreciation for the partner, making it a powerful expression of love. However, we think the last sentence is the only disappointment.

We used the same approach to evaluate many more main model generated love letters. Most of them would get an average rating of **8.0** to **10.0** points.

Discussion and Learnings

How good does the Baseline System and Main System perform?

- Both systems can produce high-quality love letters, the main model usually produces better results.
- The Main System can produce longer, more complex, and more detailed outputs.
- The Main System can create very unique and creative outputs, e.g. “I love you more than pizza”.
- The Main System requires significantly more time and resources for training than the Baseline System.
- The Main System requires more time to generate a love letter than the Baseline System.

However, there are still some problems within both systems, and some of the problems are surprising and interesting as we stated above in the Qualitative Results section.

Throughout this project, we encountered a lot of challenges, including devising a method to convert user input into prompts, and training the large language model with limited computational resources, among others. Overcoming these challenges has been a valuable learning experience and also enjoyable. We are genuinely proud of our outcomes. For a similar project, we will try to expand datasets for larger language models, and maybe come up with more plans to solve problems we have known (e.g., write additional modules and include more dictionaries).

Individual Contributions

Table 3: Individual Contribution

Hexuan Zhang	Hanxiao Wei
Collected love letters, and used the GPT-4 API to split the whole letters into sentences. Manually checked each sentence, and contributed them into the final dataset.	Collected Love Messages, and generated Love Messages from GPT-4. Wrote code to process love messages in the exact way we showed in the report’s Data Processing section.



Wrote the baseline system template. Helped to finalize the baseline, then tested and debugged the baseline system.	Progressed with developing the baseline template and created a functional system. Helped to debug and test the baseline system.
Wrote the final main system, and implemented QLoRA. Finalized, debugged, and tested the system.	Helped to test and run main models. Wrote the Transformer Module, created the words dictionaries and the Gradio UI theme.
Came up with test inputs, and the evaluation metrics.	Came up with more interesting inputs, and designed the slides/documents.
Evaluated and scored the final results.	Evaluated and scored final results.
Completed reports and slides.	Completed reports and slides.



Reference

- [1] Y. Liang and Z. Han, "Intelligent Love Letter Generator Based on GPT-2 Model," 2022 3rd International Conference on Electronic Communication and Artificial Intelligence (IWECAI), Zhuhai, China, 2022, pp. 562-567, doi: 10.1109/IWECAI55315.2022.00115.
- [2] Fagroud, F.Z., Rachdi, M., Ben Lahmar, E.H. (2022). Automatic Story Generation: Case Study of English Children's Story Generation Using GPT-2. In: Motahhir, S., Bossoufi, B. (eds) Digital Technologies and Applications. ICDTA 2022. Lecture Notes in Networks and Systems, vol 455. Springer, Cham. https://doi.org/10.1007/978-3-031-02447-4_6
- [3] Y. Belkada, T. Dettmers, A. Pagnoni, S. Gugger, S. Mangrulkar. "Making LLMs even more accessible with bitsandbytes, 4-bit quantization and QLoRA." Hugging face blog <https://huggingface.co/blog/4bit-transformers-bitsandbytes> (accessed Dec. 11, 2023)



Appendix 1: The First Sample Set of Inputs and Outputs for the Baseline Model

The prompts were automatically modified by the *Transformator* module in the main system. For the baseline system, we manually fed the following contents as the user input: ['Mike', 'You are my boyfriend', 'Since I met you in the park', 'When we were in Paris together', 'The fireworks show with you']. As a result, both models received the same prompts.

The love letter generated by the baseline model is (Text in purple are the user input):

Mike, I love you more than words can say. You are my boyfriend, my partner, my best friend, my lover, and my everything. Since I met you in the park, I knew I was in love with you, I love you more than I can express. When we were in Paris together, I felt your presence deeply, as if it were a warm embrace that comforts my soul. The fireworks show with you is the best thing that's ever happened to me.

For your convenience, the love letter generated by the main model, which is already shown in the previous Qualitative Results section, is shown here again for comparison (Texts in purple are the user input):

Mike, you are the love of my life, a man who has shown me what it means to love and be loved. I cherish you and your love. You are my boyfriend, and I love you. I don't know how you can be so cute and so handsome. Since I met you in the park, I've felt as though we were destined to be together. My heart and mind have been filled with thoughts of you. When we were in Paris together, I had the sense that we had been in love for many lifetimes, and that we were simply reminiscing about our past lives. I believe in soulmates, and you're mine. The fireworks show with you is always the best. Bye, darling.



Appendix 2: The Second Sample Set of Inputs and Outputs for Both Models

The prompts were automatically modified by the *Transformator* module in the main system. For the baseline system, we manually fed the long phases as the user input. As a result, both models received the same prompts.

The love letter generated by the baseline model is (Texts in purple are the user input):

Kitty, you are the most amazing addition to my life's journey. You are my wife, my best friend, my soulmate, and my lover, You are everything to me. You are simply beautiful and amazing, and I love you more than I can express. Since I met you in the classroom, I've found a love that's deeper than any ocean, higher than any mountain, and more profound than any words. When we were in Hawaii together, I would have never had the courage to tell you that you are the one. The W Hotels we stayed at together were the best. The Porsche we drove together was the most beautiful car I've ever seen in my life, and I'm still crazy about it. I love you, and that's a pretty big deal.

The love letter generated by the main model is (Texts in purple are the user input):

Kitty, I've found the perfect place for us, It's a secluded spot, surrounded by lush greenery, and it's the perfect place for us to spend the day together. Here's to us. You are my wife, my best friend, my partner, my motivation, and the source of all my happiness. I love you. You are simply beautiful and amazing, and I am proud to love you. And I will love you forever. Since I met you in the classroom, I have felt a special kind of love for you. You are the type of person who brings joy to others, and you make me happy. When we were in Hawaii together, I felt like the most lucky person in the world, You are such a beautiful soul. You make me want to be a better person. The W Hotels we stayed at together were the best, and I can't wait to stay in more of them. Thank you for always making me feel like royalty. The Porsche we drove together, the beautiful places we visited, and the delicious food we ate, all of these made our love stronger. My heart is forever yours. I love you. I am at rest with you, and that is enough.