

## GPT-3를 활용한 시나리오 생성 보조 시스템

조동하<sup>0</sup>, 전이슬\*, 문미경\*\*

<sup>0</sup>동서대학교 소프트웨어학과,

\*동서대학교 컴퓨터공학과,

\*\*동서대학교 소프트웨어융합대학

e-mail: jodongrami@gmail.com<sup>0</sup>, isle.jeon@gmail.com\*, mkmoon@dongseo.ac.kr\*\*

## Scenario Generation Assistance System Using GPT-3

Dongha Jo<sup>0</sup>, Isle Jeon\*, Mikyeong Moon\*\*

<sup>0</sup>Dept. of Software, Dongseo University,

\*Dept. of Computer Engineering, Dongseo University,

\*\*College of Software Convergence, Dongseo University

### ● 요약 ●

최근 자연어 처리 분야에서 언어 모델을 활용하여 문장 생성에 관한 연구가 이루어지고 있다. 기존 언어 모델을 활용하여 생성된 시나리오는 텍스트를 학습하여 활용하는 것 외에는 작가의 의도를 반영하는 것에 한계가 존재했고 문맥에 일관성 없는 모습을 보여주었다. 시나리오를 작성하는 것은 작가가 흐름을 주도하며 작업해야 하는 내용이다. 본 논문에서는 GPT-3 기반 언어 모델을 기반으로 다양한 시나리오 문장을 생성하여 작가가 선택하거나 원하는 문장을 직접 입력하는 등 작가의 의도에 부합하는 시나리오를 생성하는 보조 시스템을 제안한다. 본 연구를 통해 시나리오 생성을 포함한 문장 생성 분야의 보조 도구로 활용하여 작가의 의도를 반영하는 결과물을 생성하는 것을 목표로 한다.

**키워드:** 자연어 처리(Natural Language Processing), 시나리오(Scenario), GPT(Generative Pre-trained Transformer), 문장 생성(Text Generation)

### I. Introduction

인공지능(Artificial Intelligence)에 관한 연구는 시간이 지날수록 더욱 가속화되고 있고, 이를 활용한 시스템들이 대거 나타나고 있다. 자연어 처리는 딥러닝의 한 분야로써 많은 영역에서 활용될 수 있다. 최근에는 특정 주제로 기사나 영화의 시나리오를 작성할 수 있을 정도로 좋은 성능을 보유한 언어 모델들이 개발되고 있다 [1].

그러나 좋은 성능을 지닌 언어 모델들도 수백 단어가 넘어가는 텍스트에서는 같은 말을 반복하거나 모순되는 문장을 생성한다. 그리고 작가가 원하는 장르의 학습 데이터를 선정하는 것 외에는 시나리오에 작가의 의도를 반영하기에도 한계가 존재한다. 따라서 인공지능의 자동화라는 장점을 활용하여 시나리오를 생성할 때 학습 데이터를 선정하는 것 외 작가의 의도를 반영하는 방법이 필요하다.

### II. Related Works

LSTM(Long Short-Term Memory)을 통해 드라마 대본 생성을 진행할 때 충분한 Fine-tuning을 하지 않는다면 무한에 가까운 복잡한 토큰 배열 패턴을 학습하기에 파라미터의 양이 충분하지 않아 의미를 파악하기 어렵고 부자연스러운 문장이 생성되었다. 또한 문장 간에 모순과 전체 내용의 일치성이 떨어지는 문제가 발생하였다 [2].

본 연구에서는 자연어 처리를 위해 1750억 개의 파라미터가 학습된 GPT-3(Generative Pre-trained Transformer)를 활용한다. GPT-3는 입력되는 단어를 벡터로 전환하고 Transformer Decoder Layer에서 각 Layer마다 존재하는 18억 개의 파라미터를 통해 예측 값을 연산한다. 반환된 결과를 다시 단어로 변환하는 과정을 거친 후 조합하여 출력된다.

### III. The Proposed Scheme

Fig 1은 본 논문에서 제안하는 시스템 구조이다. 작가가 시나리오의 첫 문장을 입력하면 Input 값으로 GPT-3에 삽입되고 Transformer Decoder Layer의 예측 연산을 통해 Output를 출력한다. 본 시스템은 위 과정을 N 번 반복하여 나온 문장들과 작가의 의도를 반영하기 위해 작가가 직접 시나리오에 개입할 수 있는 입력란을 제공하여 사용자는 N+1개의 선택지를 통해 시나리오를 전개할 수 있도록 한다.

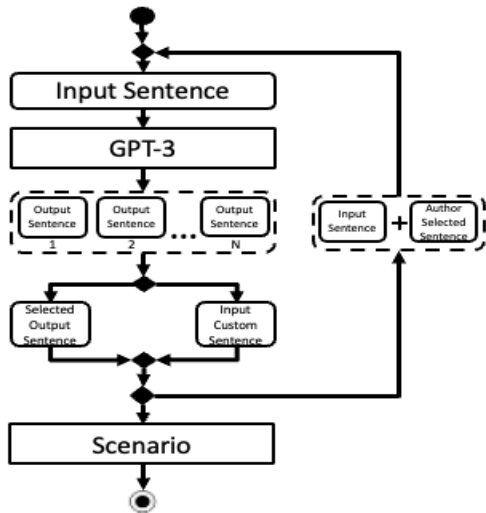


Fig. 1. System Architecture

Table 1은 Input 값에 대해 Output 값 출력 예시이다. Output 값에 모순이 없거나 작가의 의도대로 출력된 경우 제시된 선택지를 통해 시나리오를 전개해나가게 된다.

Table 1. Output 예시

Input	Once upon a time, there lived a prince in the mountains behind the Baroque kingdom.
Output #1	The prince loved the kingdom so much that he decided to become a king.
Output #2	The prince was known for his kindness, and he was beloved this people.
Output #3	He was handsome, smart, and very, very rich. Everyone loved him.

Table 2는 작가의 의도에서 벗어난 Output 값에 대해 작가가 직접 개입하여 시나리오를 전개하는 것을 보여준다.

Table 2. 작가 개입 예시

Input	Once upon a time, there lived a prince in the mountains behind the Baroque kingdom. The prince loved the kingdom so much that he decided to become a king. However, the prince did not want to take the throne away from the rightful.
Output #1	The prince decided that he would live in the mountains.
Output #2	So he built his own kingdom in the mountains.
Output #3	So he decided to live in the mountains and build his own kingdom.
Custom	So the prince decided to be recognized by the king in order to succeed to the throne.

Table 3은 본 논문에서 제안하는 시스템과 기존 시스템을 통해 출력된 예시 시나리오이며 밑줄 친 문장은 작가가 직접 입력한 문장이다. 사용자 개입 없이 생성된 기존 결과값은 문맥과 줄거리 흐름이 어색함을 알 수 있다. 본 논문에서 제안하는 시스템은 줄거리의 흐름이 문맥에서 벗어날 때 작가가 개입함으로써 전체 줄거리를 자연스럽게 만들어갈 수 있음을 알 수 있다.

Table. 3. 시스템 간 결과물 비교

제안하는 시스템	기존 시스템
<p><u>Once upon a time, there lived a prince in the mountains behind the Baroque kingdom.</u> The prince loved the kingdom so much that he decided to become a king. However, the prince did not want to take the throne away from the rightful. <u>So the prince decided to be recognized by the king in order to succeed to the throne.</u> The prince planned to defeat the enemies of the Baroque kingdom. The prince was very successful in his plan and was soon recognized by the king. The prince then took the throne and became the king of the Baroque kingdom.</p>	<p>Once upon a time, there lived a prince in the mountains behind the Baroque kingdom. The prince was very handsome, but he was also very shy. One day, he met a beautiful princess who was also shy. They fell in love and got married. They had a beautiful baby girl. The prince and princess were very happy, but they were also very poor. The prince worked hard to try to make enough money to support his family, but it was difficult. One day, the prince went into the forest to hunt for food. He saw a deer and shot it. as he was carrying the deer back to his castle, he slipped and fell down a hill. He hit his head on a rock and died.</p>

### IV. Conclusions

본 논문에서는 GPT-3를 통해 생성된 문장들을 선택해나가는 과정을 통해 작가의 의도를 반영할 수 있는 시나리오 생성 보조 시스템의 개발 내용에 관하여 기술하였다. 본 시스템을 통하여 문장 생성을 활용하는 여러 분야에 적용하여 창작의 힘을 덜어 줄 수 있을 것으로 기대한다. 향후에는 학습 언어의 97%가 한국어인 하이퍼클로바 모델과 KoGPT 모델을 활용하여 한국어로 된 시나리오를 생성하는 연구를 진행할 것이다.

## ACKNOWLEDGEMENT

본 연구는 2022년 과학기술정보통신부 및 정보통신기획평가원의 SW중심대학사업의 연구결과로 수행되었음(2019-0-01817).

## REFERENCES

- [1] <http://www.aitimes.com/news/articleView.html?idxno=132976>
- [2] Sung-Hwan Son. "Korean Drama Script Generation by using LSTM Model", July 2020.