

# KidsTale: AI-Based Children’s Story Generation System with NLP Evaluation

Hyunji Lee

Department of Information Statistics  
Chungbuk National University  
Cheongju, South Korea  
guswl1613@naver.com

Young-Seob Jeong\*

Department of Computer Engineering  
Chungbuk National University  
Cheongju, South Korea  
ysjay@chungbuk.ac.kr

October 26, 2024

## Abstract

본 연구는 AI 기반 어린이 동화 생성 시스템인 "KidsTale"에서 자연언어처리(NLP) 기법을 활용하여 동화의 장르 정확성 및 단어 적합성을 평가하는 방법을 제안한다. KidsTale은 아이가 그린 그림에서 객체를 탐지해 키워드를 추출하고, 사용자가 선택한 장르와 결합하여 사용자 맞춤형 동화를 생성한다. Chat GPT를 통해 만들어진 동화를 데이터로 하여 자연언어처리 분석을 통해 생성된 동화의 품질을 높이고, 적절하지 않은 동화는 자동으로 재생성한다. 이를 통해 어린이 사용자를 위한 개선된 사용자 경험을 제공하는 것을 목표로 한다. 이 연구에서는 감정 및 장르 분류 모델을 개발하고, 연령별 단어 난이도 사전을 구축해 Word2Vec 및 임베딩 모델을 활용한다.

**Keywords:** KidsTale, Chat GPT API, Evaluation, KoBERT, LDA, Word2Vec

## 1 Introduction

AI 기술의 발전과 함께 여러 학습 분야에서 AI 기술이 사용되고 있다. 이에 어린이 동화라는 단순한 도구에서 AI 기술을 적용한 맞춤형 동화로 확장시켜나가는 시스템을 개발하게 되었다. KidsTale은 어린이들이 그린 HTP(House-Tree-Person) Drawing Dataset으로 학습한 Yolov5로 아이들이 그린 그림에서 객체를 탐지하여 키워드를 추출하고, 모험, 추리, 공포, 우화, 사랑, 우정, 가족, 교육 8개의 장르 중 선택한 1개와 결합하여 Chat GPT API를 통해 맞춤형 동화를 생성한다. 이를 통해 아이들에게 뜻깊은 경험을 주며 독창적인 방식으로 창의력을 촉진할 수 있다. 나아가 만들어진 이야기를 본인의 방식으로 창작하는 경험을 통해 언어적 표현력과 이야기 구성 능력도 향상시켜나갈 수 있다. KidsTale에서 나온 결과물 동화를 이용해 분석을 진행하여 사용자가 선택한 장르를 바탕으로 정확한 동화를 제공하고, 이를 통해 어린이에게 감정적으로 적합한 동화를 제공하는 시스템을 향상시키고자 한다.

### 1.1 Problem

Chat APT AI 기술을 통해 생성된 동화는 감정적 적합성과 언어적 난이도 문제를 야기할 수 있다. 특히 어린이 연령대에서는 부정적인 감정이나 어려운 단어로 인해 적합하지 않은 경험을 제공할 가능성이 있다. 기존의 AI 모델들은 성인용 텍스트 분석에 초점을 맞추고 있어 어린이 대상 콘텐츠에 적합한 분석 방법이 필요하다. 따라서 부정적 감정이 과도하게 포함되거나, 어린이 연령에 맞지 않는 단어가 포함된 경우 이를 재검토하고 적절하게 수정할 필요가 있다.

### 1.2 Necessity

어린이에게 제공되는 동화는 교육적이고 감정적으로 적합해야 한다. 부적합한 콘텐츠는 아이들에게 부정적인 영향을 미칠 수 있기 때문에, 감정 분석과 장르 적합성

평가를 통해 동화의 품질을 보장하는 것이 중요하다. 이 연구는 어린이의 연령대에 맞춘 적합한 언어를 제공하고 감정적으로 안정적인 콘텐츠를 제공함으로써 맞춤형 동화 생성의 질을 향상시킬 수 있다.

### 1.3 Solution

본 연구는 YOLOv5 모델과 Chat GPT API를 사용하여 생성된 동화에 대해 KoBERT와 LDA 모델을 활용하여 감정 분석과 장르 분류를 수행한다. 또한 연령대별 단어 난이도 사전을 구축하여 각 연령대에 맞는 적절한 단어만 포함된 동화를 생성하도록 한다. 특히 감정 및 장르 분류를 통해 적합성을 평가하고 부정적이거나 적합하지 않은 콘텐츠는 자동으로 재생성하도록 설계한다.

## 2 Related Works

### 2.1 기존 제품 및 연구

기존의 동화 생성 시스템은 대부분 사용자가 직접 스토리를 작성하거나 기존 동화 데이터를 변형하는 방식에 제한된다. 이에 비해 KidsTale은 AI 기반으로 동화를 자동 생성하며 사용자의 선택에 따라 장르와 감정이 조절되는 특징을 가지고 있다. 기존 연구들은 주로 텍스트 생성과 감정 분석에 집중하고 있으나, 어린이에게 적합한 감정 분석 및 장르 분류 연구는 드물다. 특히 KoBERT와 KcBERT와 같은 모델은 한국어 텍스트 분석에 널리 사용되고 있지만, 어린이 콘텐츠에 특화된 연구는 부족한 상황이다. 이에 KidsTale은 어린이에게 감정적으로 적합한 동화를 제공하는 데 초점을 맞춘다.

**Google AI 기반 텍스트 생성 시스템:** 주로 뉴스 기사나 성인 대상 문학 작품을 생성하는 데 초점이 맞추어져 있다. 이러한 시스템은 어린이 대상 콘텐츠 생성에는 적합하지 않은 경우가 많다.

**Jinny AI:** 동화 생성 플랫폼은 AI를 사용하여 동화

를 자동으로 생성하지만 사용자가 세부적으로 스토리를 제어하거나 내용의 감정 및 난이도를 조절하는 기능이 부족하다. 따라서 사용자 맞춤형 동화를 생성하는 데 한계가 있다.

**KoBERT:** 한국어 자연어 처리 모델로 주로 한국어 텍스트의 감정 분석과 분류 작업에 사용된다. KoBERT를 사용한 대표적인 연구로는 고객 리뷰 감정 분석이 있다. 예를 들어 김준호 외(2021)의 연구에서는 한국어 고객 리뷰 데이터를 분석하여 긍정, 부정, 중립 감정을 분류하는 데 KoBERT 모델을 사용했다. 이 연구는 감정 레이블을 미리 정해진 데이터에 대해 KoBERT 모델을 Fine-Tuning하여 정확도를 크게 향상시켰다. 또 다른 연구로는 영화 리뷰 감정 분석이 있으며 이 연구에서는 리뷰 텍스트를 바탕으로 긍정적/부정적 감정을 분류하는 데 KoBERT를 사용하여 기존 모델보다 성능을 향상시켰다. KoBERT의 감정 분석 연구는 한국어 텍스트 분석에서 효과적으로 사용되고 있지만 어린이 대상 콘텐츠에 적용된 연구는 아직 부족한 상황이다.

## 2.2 차별성

KidsTale은 어린이들이 그린 그림 데이터를 바탕으로 맞춤형 동화를 생성하는 독창적인 시스템이다. 특히 감정 분석 및 장르 분류를 통해 동화의 품질을 자동 평가하며 부적절한 동화는 자동으로 수정된다. 또한 연령별 단어 난이도 사전을 구축해 각 나이에 맞는 적절한 언어를 사용하도록 한다는 점에서 기존 연구와 차별성을 가진다.

## 3 Methodology

### 3.1 Data Collection and Preprocessing

KidsTale에서 생성된 동화 텍스트 데이터는 사용자가 선택한 연령, 선택한 장르(모험, 추리, 공포, 우화, 사랑, 우정, 가족, 교육) 및 그림에서 추출된 선택 키워드 등의 메타데이터와 함께 Supabase에 자동으로 저장된다. 데이터 전처리 과정에서 텍스트 정제(불필요한 기호 제거, 띄어쓰기 수정 등)와 토큰나이징(단어를 개별 토큰으로 분리하는 작업)을 수행한다. 또한 텍스트 내에 포함된 부적절한 표현이나 어린이에게 어려운 단어를 탐지하고 수정하기 위해 단어의 난이도 사전을 구축한다. 이를 위해 국립국어원에서 제공하는 기초 어휘 목록과 AI Hub에서 제공하는 동화 줄거리 생성 데이터와 동화 이해도 테스트를 위한 질의응답쌍 생성 데이터를 활용한다.

### 3.2 Model Development

본 연구에서는 세 가지 주요 NLP 모델을 개발하여 동화 생성 시스템의 적합성과 정확성을 평가한다.

**감정 분석 모델:** 텍스트의 감정을 긍정, 중립, 부정으로 분류한다. 감정 분석은 KoBERT나 KcBERT 모델을 Fine-Tuning하여 수행되며 이 모델은 한국어 텍스트의 감정 분류에 적합한 성능을 발휘한다. 감정 분석의 목표는 아이들에게 부정적인 영향을 미치는 부정적 감정이 70% 이상인 경우 해당 동화를 다시 생성하는 것이다.

**장르 분류 모델:** 동화가 사용자가 선택한 장르에 부합하는지 평가하기 위해 주제 모델링 기법(LDA, Latent Dirichlet Allocation)을 사용한다. LDA는 텍스트를 특정 주제로 분류하는 데 사용되며 동화가 설정된 8개의 장르(모험, 추리, 공포, 우화, 사랑, 우정, 가족, 교육) 중 사용자가 선택한 장르와 얼마나 일치하는지를 평가한다. 50% 이상의 일치도를 보일 때 적합한 동화로 간주되며 그 이하일 경우 다시 생성된다.

**단어 적합성 모델:** 어린이에게 적합한 단어 사용을 보장하기 위해 Word2Vec과 같은 단어 임베딩 기법을 활용하여 텍스트 내의 어려운 단어를 식별하고 이를 더 쉬운 단어로 교체한다. 이를 위해 국립국어원의 국어학습용 말뭉치와 AI-Hub의 동화 말뭉치 데이터를 참고하여 난이도 사전을 구축한다.

### 3.3 Test and Evaluation

**성능 평가:** 동화 생성 시스템의 성능을 평가하기 위해 다양한 성능 지표를 사용한다. 혼동 행렬(Confusion Matrix)을 통해 감정 분석 및 장르 분류 모델의 성능을 측정하며 정확도(1), 정밀도(2), 재현율(3)과 같은 지표를 통해 모델의 신뢰성을 평가한다. 특히 생성된 동화가 아이들에게 감정적으로 적합하고 장르에 일관성이 있는지를 평가하는 데 중점을 둔다.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

**실사용 테스트:** 실제 어린이와 부모가 동화를 사용하도록 하여 동화를 읽었을 시에 만족도 조사를 실시한다. 피드백을 바탕으로 필요에 따라 감정 임계값과 단어 난이도 기준을 조정하여 시스템의 신뢰성을 높이고, 사용자 인터페이스와 평가 기준을 개선한다.

## 4 References

- [https://github.com/hyunjiiaiiiiiii/24\\_2\\_NLP.git](https://github.com/hyunjiiaiiiiiii/24_2_NLP.git)
- H. Kim, J. Park. "Sentiment Analysis for Korean Text Using KoBERT." Proceedings of the 2021 AI Conference, 2021.
- 김준호 외. "KoBERT 기반의 고객 리뷰 감정 분석 연구." Journal of AI and Data Science, 2021.
- S. Lee, Y. Kim. "Genre Classification with LDA: Applications in Children's Literature." Journal of NLP Research, 2020.