

## 딥러닝 기반 한국어 랩 작사 소프트웨어

\*최광희 \*\*박서진 \*\*\*권태국 \*\*\*\*구명완

서강대학교 컴퓨터공학과

\*juice500ml@gmail.com \*\*raon0211@gmail.com \*\*\*xornrbbboy@gmail.com \*\*\*\*mwkoo@sogang.ac.kr

### Software for Korean Rap Songwriting based on Deep Learning

\*Kwanghee Choi \*\*Sojin Park \*\*\*Taeguk Kwon \*\*\*\*Myoung-Wan Koo

Department of Computer Science & Engineering, Sogang University

#### 요약

해당 소프트웨어는 딥러닝 기반의 언어 모델을 이용하여 한국어 랩을 더 효과적으로 작성할 수 있도록 돕는다. 이 소프트웨어는 단순히 가사를 생성하는 데에 그치지 않고, 라임을 맞추고자 하는 대상 단어와 앞뒤 문맥이 주어졌을 때, 라임과 맥락에 맞는 단어 목록을 추천한다. 작사가는 사용자 친화적인 문서 편집 인터페이스를 통하여 언어 모델과 적극적으로 상호작용을 해 나가며 효율적으로 가사를 만들어나갈 수 있다.

#### 1. 서론

##### 1.1 작품의 동기

인간의 삶에 있어 음악은 매우 중요한 위치를 가진다. 그 사실을 증명하듯, 컴퓨터 기반의 작곡 프로그램인 디지털 음성 워크스테이션(Digital Audio Workstation) 시장은 해가 지날수록 성장하고 있다[1]. 그에 반해 작사 소프트웨어 시장은 사실상 전무한 상태이다. 인공지능이 폭발적으로 성장하기 전 소프트웨어는 창의성이 필요한 작업을 하기에 한계가 있었기 때문이다. 하지만, 비지도 학습(Unsupervised learning) 기반의 대형 언어 모델(Large-scale language model)은 그 편견을 부수고 있다[2, 3]. 이미 구글의 번역 서비스는 전세계인이 사용하고 있고[4], 챗봇(Chatbot)은 이미 일상에 깊게 녹아들었다[6].

인공지능 기술을 이용한 소프트웨어는 인간의 작업 능력을 향상시키고 있다. 구글의 번역 서비스는 번역가들에게 초벌번역 서비스를 제공함으로써 작업 속도를 2-30% 가량 향상시켰다고 알려져 있다[6]. 지금부터 소개할 한국어 랩 작사 소프트웨어는 구글 번역 서비스와 같이 언어 모델을 활용하여 작사가가 보다 손쉽게 노래 가사를 쓸 수 있도록 도와주는 것을 목표로 한다.

작사 과정의 생산성을 높이기 위해서는 작사가가 작곡 소프트웨어와 어떻게 같이 협업하고 상호작용할지를 설계하는 것도 매우 중요한 위치를 가진다. 고로, 이 보고서는 소프트웨어와 작사가의 상호작용을 고려한 작사의 정의를 먼저 살펴보고, 언어 모델을 어떻게 활용할 수 있을지를 살펴본 뒤, 실제 어플리케이션 화면 캡처를 통해 어떤 식으로 상

호작용을 설계했는지 소개한다.

#### 2. 문제 정의

이 소프트웨어는 처음부터 모든 곡을 작사하고자 하지 않는다. 그 대신, 일부가 완성된 곡이 있을 때 나머지 부분을 채워넣고자 한다. 랩과 같은 곡 형태에서 작사는 더 어려운 일이 되는데, 라임(Rhyme)을 맞춰야 하기 때문이다[7]. 라임이란 일부 곡 형태에서 필수 불가결한 요소이자 약속 체계이며, 일반적으로 한국어에서는 중성 모음을 비슷하게 뚫으로써 실현된다[7]. 라임을 통해 가사 자체에 리듬감을 줄 수 있으며, 그를 통해 음악적 가치를 더욱 배가하게 된다[8].

이 소프트웨어는 작사의 정의를 좁혀 한국어 랩 가사를 작성하는 데에 도움을 주고자 한다. 일부 가사가 주어지고 라임을 맞추어야 하는 대상 단어가 주어졌을 때, 가사가 될 수 있는 여러 후보 단어들을 추천하는 문제에 집중한다.

#### 3. 기존 연구 및 소프트웨어

언어 모델이 급격하게 발전하며 모델을 기반으로 작사를 하거나[9,10] 시를 쓰고자 하는 시도들이 있어 왔다[11,12].

[9]의 경우 영어 랩을 처음부터 자동으로 생성할 수 있는 언어 모델을 학습하는 데에 집중하면서, 라임에 대한 정량적 평가를 수행하기 위해 라임 밀도(Rhyme density measure)라고 하는 개념을 소개한다.

이 소프트웨어는 [9]의 라임 밀도를 응용하여 한국어 가사 밀도를 정의한다.

[10]은 한국어 랩을 생성하는 것을 목표로 한다. [10]과 이 보고서에서의 작사의 정의는 매우 비슷하다. 그러나 [10]은 가사의 맨 처음이나 마지막에 들어갈 단어만을 추천해주지만, 이 소프트웨어는 어디에 위치한 단어도 추천할 수 있다는 개선점을 가지고 있다.

[12]는 각 작가가 남긴 시(운문) 데이터를 바탕으로 다음 단어를 생성한다. 시와 랩은 라임을 맞추어야 한다는 점에서 비슷한 속성을 가지고 있으나, 문제의 범주가 다르다.

[13]는 특정 단어와 라임이 맞는 단어를 검색할 수 있는 소프트웨어이지만, 앞뒤에 어떤 가사가 들어갈지에 대해 판단하지 않는다. [14]도 마찬가지로 모음만 입력하면 단어만 검색되는 기능을 가지고 있다. 이 보고서는 문맥에 어울리면서 동시에 라임이 맞는 추천을 함으로써, 문맥을 고려하지 않는 [13,14]을 개선한다.

#### 4. 자동 인코딩 기반 언어 모델

최근의 비지도 학습 기반 대형 언어 모델은 언어 코퍼스(Language corpus)을 이용하여 학습하고, 흔히 두 가지 학습 방법 중 하나를 채택한다. 자동 회귀 (Autoregressive) [2]와 자동 인코딩 (Auto-encoding) [3]이 그것이다. 그 중 자동 인코딩 기반 언어 모델은 코퍼스로부터 임의의 문장을 선택하고, 해당 문장 내의 임의 단어를 가린 후(Masking), 가려진 단어 앞뒤의 문장을 보고 원래 단어를 다시 복원해내는 문제(Task)를 해결한다. 과정에서 일반적으로 활용할 수 있는 언어 모델(General language model)을 학습한다.

자동 인코딩 언어 모델의 학습 방식은 라임이라는 특징을 제외한다면 위 작사의 정의에서 다른 문제 정의와 동일함을 알 수 있다. 이 소프트웨어는 해당 방식으로 학습시킨 bert-kor-base[15]를 활용하여 추천 가사를 생성한다. 그리고 라임에 알맞은 단어만 걸러내어 적당한 가사를 추천한다.

#### 5. 한국어 라임 밀도

한국어 랩에서 라임을 맞추기 위한 방법 중 하나는 비슷한 중성 모음을 활용하는 것이다. 해당 전략은 유명 래퍼 "피타입"의 각운 노트[16]에서도 확인할 수 있다. 예컨대 "ㅏ"와 "ㅑ"는 비슷한 소리 특성을 갖기 때문에 비슷한 중성 모음이라고 할 수 있다[16]. 영어에서도 비슷한 방식으로 라임을 맞출 수 있다. 영어 랩 가사를 생성하는 [9]의 경우, 모음이 얼마나 비슷한가를 라임 밀도로 정의하여 사용한다. 위와 같은 사실을 활용해서 한국어 랩 가사에서 얼마나 라임이 잘 맞는지에 대한 평가 지표를 설계해볼 수 있다.

초성	중성	종성
ㄱ ㄷ ㅌ ㅍ ㅊ ㅋ ㅅ ㅈ ㅊ ㅌ ㅍ ㅊ ㅋ	ㅏ ㅑ	ㄴ ㄹ ㄷ
ㄴ ㄹ ㄷ	ㅓ	ㄹ
ㅏ ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ	ㅕ ㅛ ㅜ ㅠ	ㄱ ㅋ ㄷ ㅌ ㅍ ㅊ ㅊ ㅌ ㅍ ㅊ ㅋ
ㅇ ㅎ	ㅕ ㅛ	ㅏ ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ ㅕ ㅛ ㅜ ㅠ ㅕ ㅛ ㅜ ㅠ
	ㅕ ㅛ	
	ㅕ ㅛ ㅕ ㅛ ㅕ ㅛ ㅕ ㅛ	
	ㅕ ㅛ	

그림 1 피타입의 각운 노트

먼저 [9]와 비슷하게 두 가사에서 모음만 모두 추출한다. 모음만 남은 두 가사에 대해 LCS(Longest common subsequence) 알고리즘을 적용시키면 두 가사에 얼마나 많은 모음이 겹쳐있는지 셀 수 있다. 두 모음만 남은 가사 중 더 긴 쪽의 길이로 겹친 모음의 갯수를 나누면, 0과 1 사이의 정규화된 점수(Normalized score)로 나타낼 수 있다. 여기서 [16]을 활용하여 모음이 완전히 동일하지 않더라도 비슷한 모음이면 동일한 모음으로 정의하면 한국어 가사에서의 라임 밀도를 정의할 수 있다.

bert-kor-base[15]를 통해 생성한 추천 가사들 중, 해당 정의를 통하여 라임을 따르지 않는 가사들을 배제할 수 있다. 여기서 작사가는 하여금 얼마나 라임을 엄밀하게 지키는지에 대한 문턱 값(threshold)을 정의하여 라임에 대한 엄밀성을 높이거나 줄일 수 있다(controllable). 이 소프트웨어에서는 해당 문턱 값을 가장 처음에 단 한번 설정하는 것이 아니라, 언제든지 수정할 수 있다. 이로써 작사가와 소프트웨어는 더욱 효과적으로 상호작용할 수 있다.

#### 6. 작사 인터페이스(Interface)

이 논문에서의 좁은 작사의 정의를 다시 살펴보도록 하자. 작사자에게 소프트웨어가 새로운 가사를 추천해주기 위해서는 라임을 맞춰야 하는 어떤 특정 단어가 존재하고, 미완성된 가사가 주어지며, 어떤 위치에 단어가 추가되어야 할지에 대한 정보가 필요하다. 여기서 bert-kor-base[15]를 통해 미완성된 가사에서 문맥을 고려한 추천 단어를 생성하고, 생성한 단어와 라임의 특정 단어의 한국어 라임 밀도를 측정한다. 이로써 라임이 맞는 단어만 추천되도록 한다.

최종 추천 단어는 단 한 개가 아닌 가능한 한 많은 단어를 보여줌으로써 궁극적으로 작사가가 선택할 수 있는 선택지를 여럿 제공한다. 해당 기능에 대해 작사가와 효과적으로 상호작용할 수 있도록 실제로 구현된 소프트웨어 UI(User interface)를 소개한다.

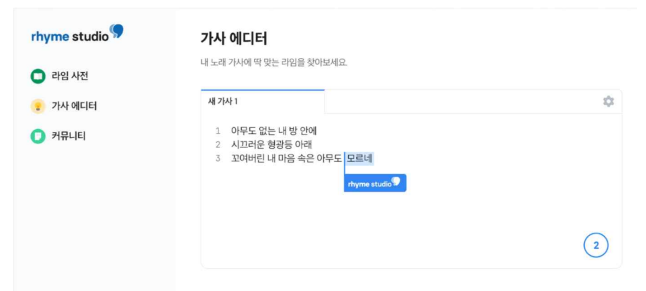


그림 2 사용자가 충분한 길이의 가사를 입력했을 때, 소프트웨어가 가장 적합한 다음 단어를 추천하는 화면 캡처

먼저 작사가는 소프트웨어의 "가사 에디터"에 접속하고, 쓰고자 하는 노래 가사의 초안을 작성한다. 시스템은 충분한 길이를 가지는 가사가 입력되면, 전후 맥락과 라임 조건에 가장 맞는 다음 단어를 추천한다.



그림 3 사용자가 추가로 다양한 추천 단어를 조회하는 화면

작사가는 처음 보이는 단어 이외에도 다양한 추천 단어를 조회할 수 있다. 추천 단어 목록의 오른쪽에서는 각 단어가 전체 가사의 맥락에 맞는 정도와 라임의 일치율(라임 밀도)이 시각화된 형태로 나타난다. 추천 단어는 합산 점수가 높은 순으로 표시된다.



그림 4 사용자가 작성한 가사에 대해서 추천 단어를 제시하는 화면

작사가는 이미 작성한 가사에 대해서도 바꾸어 쓸 수 있는 새로운 선택지를 발견할 수 있다. 가사의 각 부분에서 바꾸어 쓸 수 있는 단어가 있다면 파란색으로 밑줄이 그어진다. 작사가는 그 단어를 클릭함으로써 의미와 라임이 비슷한 다른 낱말을 확인할 수 있다.



그림 5 이미 작성된 가사에 대해서 바꾸어 쓸 수 있는 단어 목록을 확인하는 화면

작사가가 밑줄이 그어진 단어를 클릭하면 바꾸어 쓸 수 있는 단어의 추천 목록을 확인할 수 있다. 여기에서도 낱말의 목록은 가장 노래의 분위기와 라임에 맞는 순으로 표시된다.

## 7. 결론

이 보고서는 그동안 작곡에 비해 주목을 받지 못했던 작사 소프트웨어 시장을 조명하면서, 시장에서 하나의 플레이어가 될 수 있는 "한국어 랩 가사 생성 소프트웨어"의 설계와 그 구현을 소개한다. 이 소프트웨어는 단순히 가사를 생성하는 것에 그치지 않고 작사가와 소프트웨어가 서로 적극적으로 상호작용하면서 완결된 가사를 만들어갈 수 있도록 한다. 먼저, 해당 보고서에서 살펴보는 좁은 정의의 작사에 대해 살펴보고, 해

당 문제 정의에 대해 기존에 존재하는 연구들과 소프트웨어들을 살펴본다. 그 후, 후보 가사를 생성하기 위해서 최근 들어 뛰어난 성능을 보이는 자연어 모델 중 자동 인코딩 모델을 활용해야 하는 이유를 살펴본다. 추천된 후보 가사들 중 라임에 맞는 가사들만 추려내기 위해 한국어 라임 밀도를 정의하고, 실제로 구현된 소프트웨어 화면 캡처를 통하여 어떠한 형식으로 작사가와 소프트웨어가 상호작용하는지 살펴본다.

## 참고 문헌

- [1] Verified Market Research, "Global Digital Audio Workstation Market Overview," Global Digital Audio Workstation Market Size By Type, By Component, By End-User, By Geographic Scope And Forecast, 2021.
- [2] A. Radford, et al., "Language models are unsupervised multitask learners," OpenAI blog 1.8 (9), 2018.
- [3] J. Devlin, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding," Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2019.
- [4] 김범수, 구글 번역 최고담당자의 예상밖 답변 "번역기가 인간을 완전 대체하는 시점은 오지 않을 수도", 조선비즈, 2017.
- [5] 카차 한계, "자동 번역: 점점 개선되고 있으나 아직 완벽하지 못하다", 독일문화원(괴테-인스티튜트), 2012.
- [6] Luo, Xueming, et al. "Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases." Marketing Science 38.6, 2019.
- [7] 김성훈, "랩 창작 방법: 라임과 플로우를 중심으로," 석사학위논문, 숭실대학교, 2015.
- [8] P. Edwards, "How to rap," Chicago Review Press, 2009.
- [9] M. Eric, et al., "Dopelearning: A computational approach to rap lyrics generation," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016.
- [10] 손성환 et al., "딥러닝 기법을 이용한 노래 가사 생성 시스템," 제 30 회 한글 및 한국어 정보처리 학술대회, 2018.
- [11] A. Zugarini, et al., "Neural Poetry: Learning to Generate Poems using Syllables," International Conference on Artificial Neural Networks, 2019.
- [12] M. C. Santillan, et al., "Poem Generation using Transformers and Doc2Vec Embeddings," Intl. Joint Conference on Neural Networks, 2020.
- [13] "RHYMENOTE," <http://rhymenote.com/list>, 2021년 11월 6일 접속.
- [14] "국어사전," 네이버, 2021년 11월 6일 접속.
- [15] 김기영, "Pretrained Language Models For Korean," Github ([github.com/kiyoungkim1/LMkor](https://github.com/kiyoungkim1/LMkor)), 2021년 11월 6일 접속.
- [16] "피타입이 직접 설명하는 각운 노트," <https://music.naver.com/promotion/specialContent.nhn?articleId=4102>, 2015년 1월 1일 접속.