

딥러닝을 이용한 BGM 음원 작곡 서비스 설계 및 구현

김영훈*, 윤성열*, 김병우*, 신현우*, 황규영**, 한연희***†

*한국기술교육대학교 컴퓨터공학부

**한국기술교육대학교 대학원 컴퓨터공학과

***한국기술교육대학교 첨단기술연구소

e-mail: {kyh9774, tjdduf3710, bwmichael, fun5733, to6289, yhhhan}@koreatech.ac.kr

Design and Implementation of BGM Composition Service Using Deep Learning

Young-Hoon Kim*, Sung-Yeol Yoon*, Byung-Woo Kim*, Hyun-Woo Shin*,
Gyu-Young Hwang**, Youn-Hee Han***

*Dept of Computer Science and Engineering, KoreaTech University

**Dept of Computer Science and Engineering, Graduate School of KoreaTech University

***Advanced Technology Research Center, KoreaTech University

요 약

인공지능에 대한 연구가 계속되고 있지만 음악, 미술, 문학 등 창의력과 예술성을 요구하는 부문에서는 인공지능을 적용하기 어렵다. 하지만 그중 음악 부문에 인공지능을 접목시켜 기존에 없던 곡을 작곡해보고자 한다. LSTM 기반의 모델을 ABC Notation 악보 데이터를 활용하여 학습시켜 사용자들이 음악적 지식 없이도 새로운 음악을 작곡할 수 있도록 딥러닝을 기반으로 한 BGM 음원 작곡 서비스를 제안한다.

1. 서론

최근 인공지능은 우리 삶에 큰 영향을 끼치고 있으며 인간이 하던 일들의 대체재로 부상하고 있다. 그중 음악, 미술, 문학 등 창의력을 요구하는 예술성이 요구되는 분야는 미래에 인공지능이 대체할 수 없다는 것이 지배적인 의견이다. 하지만 최근 반 고흐 화풍의 그림을 그리는 예술적인 분야의 인공지능이 개발되는 것을 보면 이런 의견에 의문이 생길 수밖에 없다.

최근 뉴미디어로 주목받고 있는 개인 인터넷방송은 개인 혼자서 방송의 모든 콘텐츠를 제작하지만, 현실적으로 모든 콘텐츠를 개인 혼자서 제작하는 것은 어렵다. 따라서 썸네일(Thumbnail), BGM(Background Music), 인트로(Intro), 아웃트로(Outro) 등의 콘텐츠는 외주를 맡기거나 저작권이 없는 콘텐츠를 사용한다. 그중 BGM(Background Music)은 대부분 개인 인터넷 방송인들의 경우 저

작권이 없는 음악들을 가져와 사용한다. 하지만 저작권에 저촉되는 음악을 사용할 경우 콘텐츠 자체의 수익창출이 어려워진다.

인공지능이 발전함에 따라 사람보다 뛰어난 인공지능 바둑 프로그램인 알파고와 등장하게 되면서 딥러닝을 활용하여 다양한 인공지능 프로그램들이 개발되고 있다. 이러한 인공지능을 사용하여 저작권이 없는 음악을 대량으로 생산하면 많은 게임 개발자나 방송인에게 값싸게 혹은 무료로 배포할 수 있다. 기존에 [1, 2]와 같은 작곡 관련 연구가 이를 나타낸다. 뛰어난 작곡가들은 이미 많이 존재하지만, 현시대의 흐름에 맞추어 본 연구에서는 어떻게 하면 인공지능이 곡을 스스로 학습하여 창작의 영역까지 해낼 수 있을지를 연구하고 이를 해결할 수 있는 솔루션 개발을 목표로 한다. 본 연구에서 해결해야 할 현상은 ‘인공지능이 곡을 스스로 학습하고 창작을 할 수 있어야 한다.’로 요약할 수 있다. 만약에 인공

† 교신저자: 한연희(한국기술교육대학교)

이 논문은 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No.2018R1A6A03025526 및 No.2016R1D1A3B03933355)

지능이 곡을 만들면 저작권이 없어지게 되고 게임 개발자나 방송을 하는 사람에게는 공짜로 곡을 사용할 수 있기에 비용적인 측면을 절약할 수 있다. 또한, 작곡가가 곡을 만드는 시간을 줄여주어 시간적인 측면에서도 절약할 수 있다. 따라서 인공지능이 딥러닝을 통해 스스로 학습하여 곡을 창작하는 것을 본 연구에서 제안한다.

결과적으로 본 논문에서 제안하는 시스템은 딥러닝(Deep Learning)을 이용하여 새로운 곡을 작곡해주고 개인 인터넷 방송인이 저작권 없는 음악을 찾거나 외주를 맡겨 작곡할 때 소비하는 시간을 줄여주며 저작권법에 저촉되는 음악을 사용하여 콘텐츠의 수익창출에 어려움이 없도록 도움을 준다.

2. 작곡을 위한 딥러닝 학습 방법

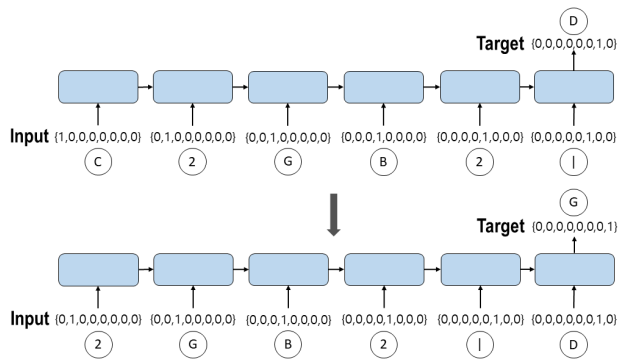
본 연구는 ABC Notation 악보를 데이터로 사용하고 있다. 이 악보 데이터를 분위기에 따라 분류한 후 데이터들을 전처리하여 텍스트를 인덱스로 매핑한다.

ABC Notation이란 악보를 A부터 G까지 문자 표기법을 사용하여 나타낸 것이다. 여기에는 번호, 제목, 작곡가, 음표, 길이, 음계 등이 포함되어 있고 각 알파벳 별로 의미하는 것이 정해져 있다. 이 표기법은 계속해서 최신화가 되어가고 있다. <그림 1>을 보면 각 알파벳이 무엇을 의미하는지 알 수 있다 [3].

Field name	Examples and notes
A:area	A:Donegal, A:Bampton (<i>deprecated</i>)
B:book	B:O'Neill's
C:composer	C:Robert Jones, C:Trad.
D:discography	D:Chieftains IV
F:file url	F:http://a.b.c/file.abc
G:group	G:flute
H:history	H:The story behind this tune ...
I:instruction	I:papersize A4, I:newpage
K:key	K:G, K:Dm, K:AMix
L:unit note length	L:1/4, L:1/8
M:meter	M:3/4, M:4/4
m:macro	m: ~G2 = {A}G{F}G
N:notes	N:see also O'Neill's - 234
O:origin	O:UK; Yorkshire; Bradford
P:parts	P:A, P:ABAC, P:(A2B)3
Q:tempo	Q:"allegro" 1/4=120
R:rhythm	R:R, R:reel
r:remark	r:I love abc
S:source	S:collected in Brittany
s:symbol line	s: lpp! ** lfl
T:tune title	T:Paddy O'Rafferty
U:user defined	U: T = !trill!
V:voice	V:4 clef=bass
W:words	W:lyrics printed after the end of the tune
w:words	w:lyrics printed aligned with the notes of a tune
X:reference number	X:1, X:2
Z:transcription	Z:John Smith, <j.s@mail.com>

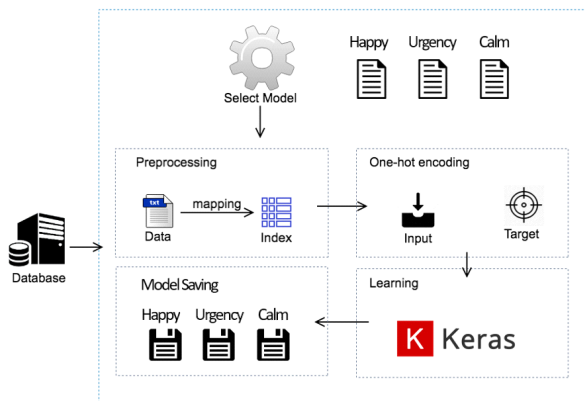
<그림 1> ABC Notation 표기표

위와 같은 ABC Notation이 Text 형태의 데이터로 저장이 되면 전처리 과정에 들어가게 된다. 이 과정을 거치게 되면 곡에 있는 모든 음들은 사전 형태로 매핑이 된다. 이를 입력데이터로 사용하기 위해 벡터 형태로 바꿔주게 되는데 이때 사용하는 것이 원 핫 인코딩(One-hot encoding)이다. 입력데이터가 설정되면 LSTM으로 학습이 이루어지게 되는데 여기서 입력데이터의 개수는 6개로 설정한다[4]. 너무 작으면 “”과 같이 여담음이 있는 데이터들에서 오류가 발생할 수 있고 너무 길게 되면 예측이 제대로 되지 않는 현상이 발생할 수 있다. 따라서 본 논문은 스택을 6으로 설정하여 문자들의 오류를 줄이고자 한다. 6개로 이루어진 하나의 데이터 셋과 이 데이터 셋 다음에 오는 하나의 데이터를 타겟 데이터로 하여 학습이 된다. 아래 그림과 같이 한 스택씩 밀리면서 학습 과정이 반복된다.



<그림 2> 학습 과정

위 학습 과정은 Epoch에 따라 반복되고 학습이 완료되면 모델이 생성된다. 학습 시스템 구성도는 다음과 같다.



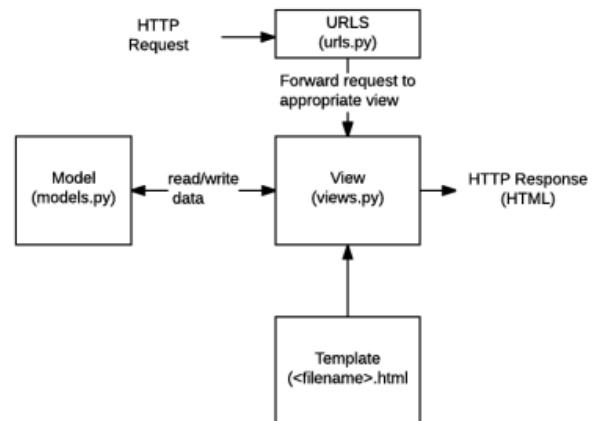
<그림 3> 심층학습 시스템 구성도

결과적으로 Preprocessing에서 텍스트 데이터를 인덱스로 매핑하고 이를 원 핫 인코딩으로 벡터화하여 입력데이터의 기준에 맞게 데이터를 변환한다. 이후 Learning 단계에서 LSTM을 통해 데이터를 학습하여 모델을 생성한다. 마지막으로 Model Saving에서 생성된 모델을 사용자가 초기에 설정한 임의의 곡 분위기로 저장한다.

3. 서비스 시스템 설계

Epoch에 따라 생성된 모델들을 Django로 작성된 웹서버에 저장하면 사용자들은 원하는 분위기를 웹에서 선택할 수 있게 되고 그에 맞는 새로운 곡이 작곡된다. Django의 진행 과정은 다음과 같다. urls.py에 저장된 URL mapper가 요청 URL을 기준으로 HTTP 요청을 적절한 view로 보내주기 위해 사용된다. 단일 함수를 통해 모든 URL 요청을 처리하는 것은 가능하지만 분리된 뷰 함수를 작성하는 것이 각각의 리소스를 유지·보수하기 쉽다. 만약 지정된 URL 패턴과 일치하는 HTTP 요청이 수신된다면 views.py 안에 저장된 관련 view 함수가 HTTP

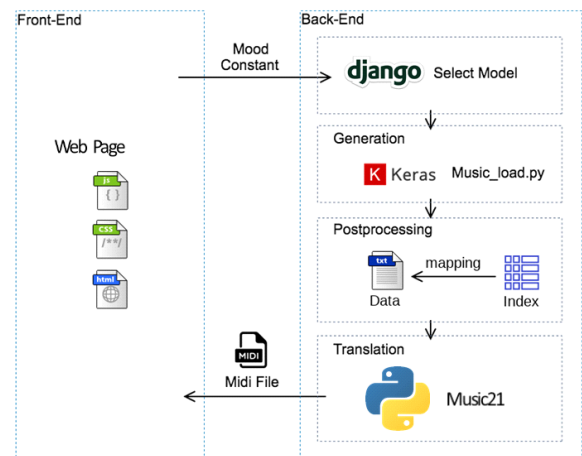
응답을 반환하게 된다. Django의 기본적인 흐름도는 <그림 4>와 같다 [5].



<그림 4> Django 흐름도

작곡이 완료된 곡은 Music21 라이브러리를 통해 Midi 파일로 변환된다. 사용자가 재생을 원할 경우, 플레이어로 재생이 되고 다운로드를 원할 경우, Midi 파일로 다운로드가 진행된다.

서비스 시스템의 구성도는 다음과 같다.



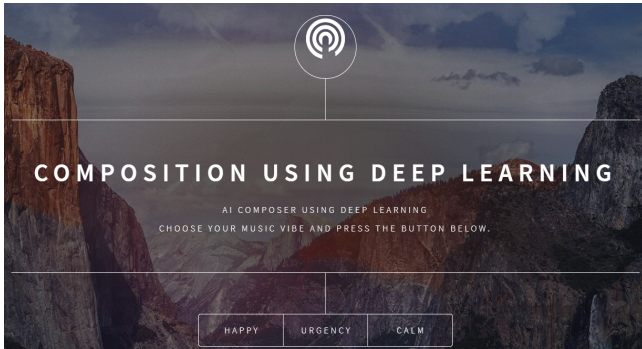
<그림 5> 서비스 시스템 구성도

Web Page는 사용자가 시스템의 기능을 이용하기 쉽도록 웹사이트의 형태로 서비스를 제공하고 Generation은 저장된 모델을 통해 새로운 곡을 작곡하는 과정이다. Postprocessing은 Generation을 통해 나온 출력 데이터를 다시 텍스트로 매핑하여 Midi 파일 변환에 맞는 데이터에 맞는 형태로 만든다. Translation은 Postprocessing을 통해 나온 ABC Notation을 Music21을 이용하여 Midi 파일로 변환한다.

4. 서비스 시스템 구현 결과

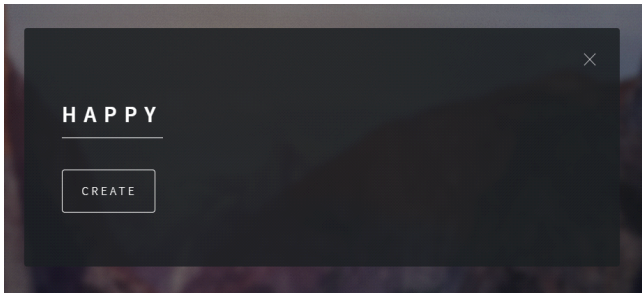
본 서비스의 실행 웹 페이지는 다음과 같다. 웹페

이지에 접속하게 되면 <그림 6-1>과 같은 페이지가 열린다.



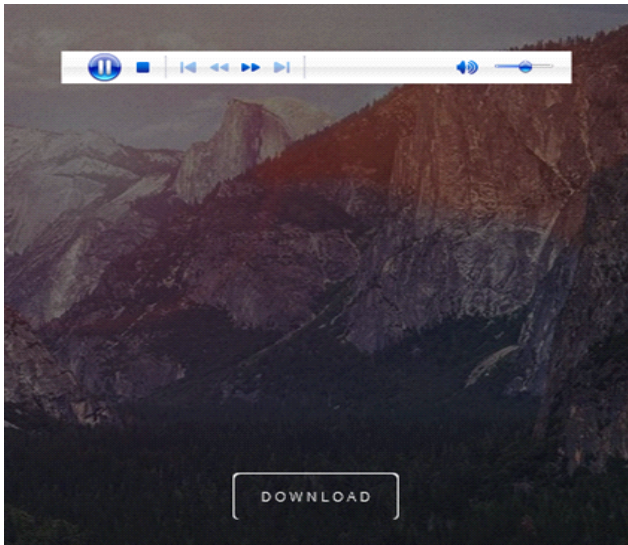
<그림 6-1> 시작 페이지

사용자는 웹페이지 중간에 Happy(밝음), Urgency(긴박), Calm(잔잔) 버튼이 있는 것을 확인할 수 있다. 자신이 작곡하고 싶은 분위기의 버튼을 클릭하게 되면 <그림 6-2>와 같은 새로운 화면이 열리게 된다.



<그림 6-2> 분위기 선택 후 페이지

다른 분위기로 전환하기를 원하면 우측 상단에 있는 x 표시를 누르고 Midi 서비스를 이용하려면 Create버튼을 누르면 된다. 버튼을 누르게 되면 <그림 6-3>과 같은 화면이 열리는 것을 확인할 수 있다.



<그림 6-3> 다운로드 페이지

사용자는 플레이어로 작곡된 곡을 들을 수 있고 다운로드를 원하는 경우 화면 하단에 있는 'DOWNLOAD' 버튼을 눌러 다운로드를 받을 수 있다.

5. 결론

본 논문은 ABC Notation을 분위기별로 분류하여 딥러닝으로 학습하고 여기서 만들어진 모델을 통해 기존에 없던 새로운 음악을 사용자가 이용할 수 있다. 이 사용자들은 음악적 지식 없이 작곡할 수 있고 인공지능으로 작곡을 했기 때문에 기존의 작곡가가 만드는 곡보다 시간·경제적인 소요를 줄일 수 있다. 여기서 키포인트는 시간·경제적인 소요이다. 기본적으로 인공지능으로 작곡하는 곡을 배포할 때 비용적인 측면을 고려해야 한다. 또한, 작곡가가 만드는 곡보다 질적으로 너무 떨어지면 안 된다. 향후 본 논문에서 제안한 학습 방법을 고도화하여 높은 질의 음악을 작곡하는 방법을 연구할 예정이다.

참고문헌

- [1] Saurabh Malgaonkar, Yudhajit Biswajit Nag, Rohit Devadiga, Tejas Hirave, "An AI based intelligent music composing algorithm: Concord," International Conference on Advances in Technology and Engineering (ICATE), 2013.
- [2] Omar Lopez-Rincon, Oleg Starostenko, Gerardo Ayala-San Martín, "Algorithmic music composition based on artificial intelligence: A survey," International Conference on Electronics, Communications and Computers (CONIELECOMP), 2018.
- [3] ABC Notation <http://abcnotation.com/>
- [4] Shuai Li, Qi Wang, Xiaojie Liu, Jienan Chen, "Low Cost LSTM Implementation based on Stochastic Computing for Channel State Information Prediction," IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), 2018.
- [5] 김석훈, 파이썬 웹 프로그래밍, 한빛미디어, 2018