

# 딥러닝을 이용한 낭만파 클래식 음악 창작 방법론

LSTM & Attention

2020. 12. 18

지성태

# 1.Intro

## ● AI 음악 창작

- **음악** : 음정, 박자, 화성, 셈여림 등 많은 변수가 복합된 창작물
- 이미지, 텍스트에 비해 연구, 활용 사례 많지 않음
- 단순 멜로디 창작이나 대중음악 장르 일부 진전 있음

딥러닝 기법을 이용,

클래식 곡 중 가장 복잡하고 긴 **교향곡 생성** 도전,  
이에 필요한 **기법 연구**, 보다 나은 결과물 위한 방법론 제시

## ● Difficulty

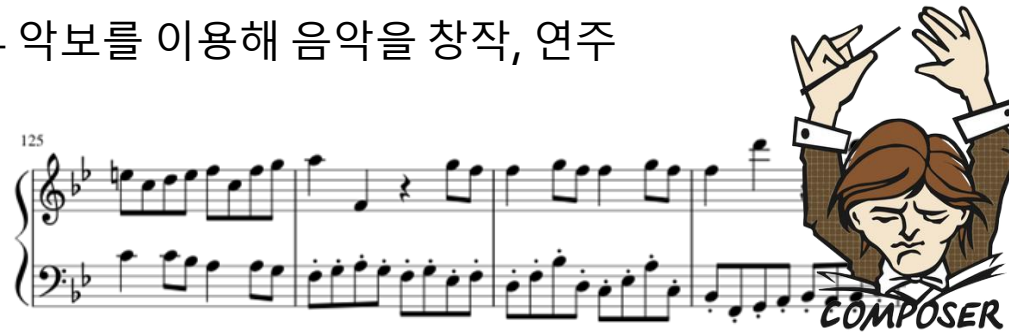
- 많은 경우의 수 / Sparse한 데이터 구조
- 예측 모델의 정확도(Accuracy)가 음악적 완성도(Quality)를 의미하지는 않음
- 결과물을 눈으로 쉽게 확인할 수 없음(Intangible)

# 1.Intro

## Way of Music Generating

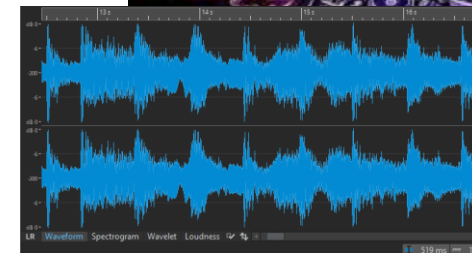
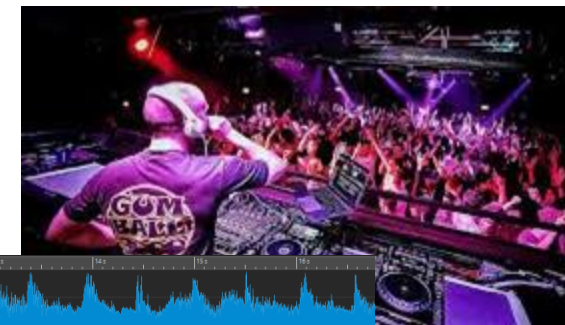
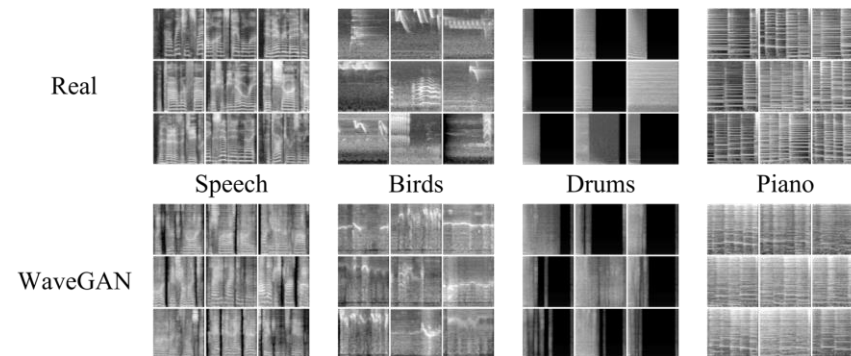
### ● Notation (Symbolic Music)

- 악보를 이용해 음악을 창작, 연주



### ● Audio (Timbre) Synthesis

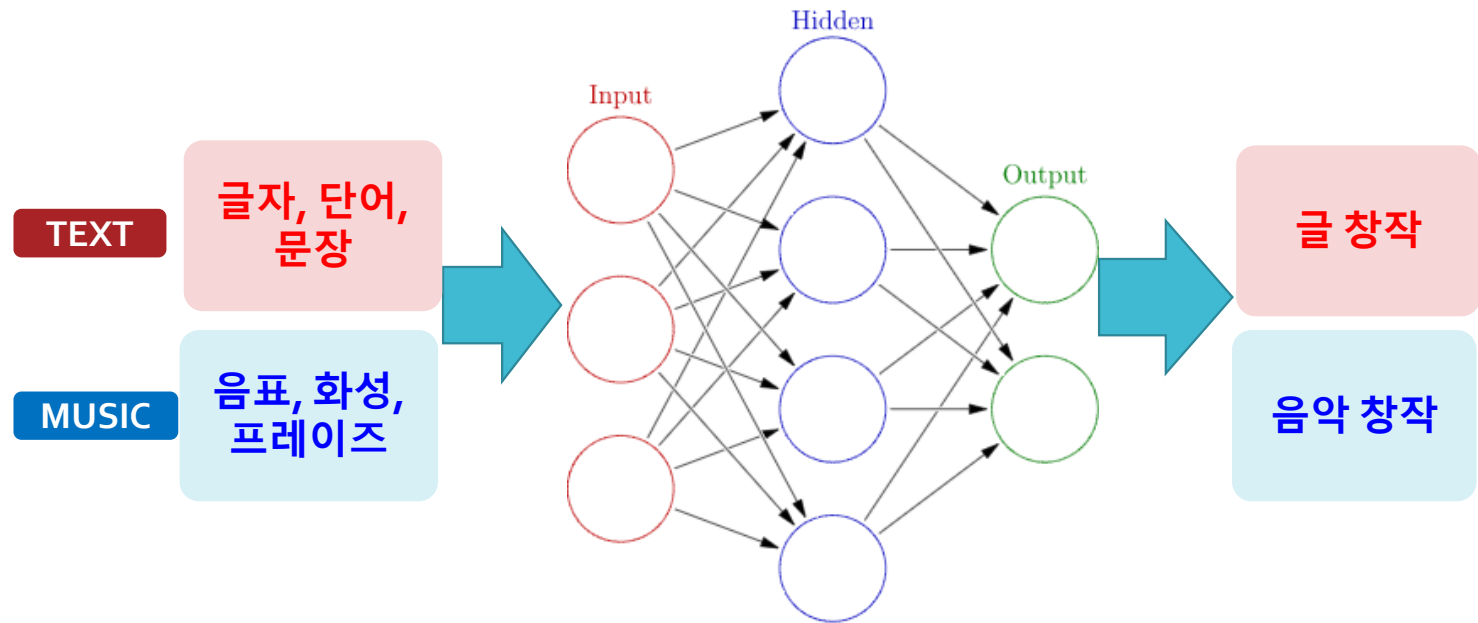
- 귀에 들리는 오디오 그대로 이용(합성), 창작
- CNN + GAN



## 2. Concept

“음악은 인류의 만국 공통어”  
- Henry W. Longfellow, John Wilson

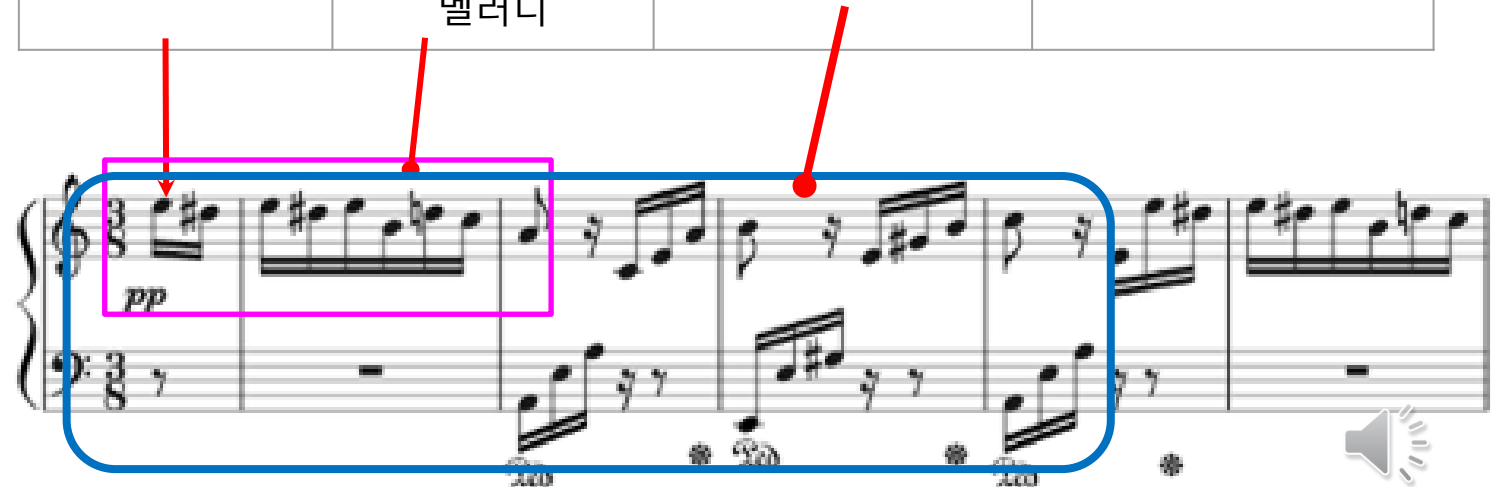
- AI 음악 생성에 자연어 처리 방식 적용 추세 [CNN(X) RNN(O)]



## 2. Concept

“내가 그의 이름을 불러주기 전에는 그는 다만 하나의 몸짓에 지나지 않았다.  
내가 그의 이름을 불러주었을 때, 그는 나에게로 와서 꽃이 되었다.”

글자(단어)	구	문장	문단
내가, 그의, 꽃, 되었다...	잊혀지지 않는, 하나의 몸짓, ...	그는 다만 하나의 몸짓에 지나지 않았다	내가 그의 이름을 불러주기 전에는 그는 다만 하나의 몸짓에 지나지 않았다.
음표(Note)	모티브 (동기)	작은 악절	큰 악절
길이, 음정, 화음	1마디 내외 멜러디	4-8마디 단위	16 또는 32마디



### 3. 연구 사례

- LSTM (2016~)

- 사건 사이에 알 수 없는 크기 및 시간에 따른 정보를 분류, 처리 및 예측하는 RNN의 일종
- 시계열 데이터나 연속적인 결과물 모델링에 주로 쓰임

- Google Magenta Project

- Attention 기반 Music Transformer (2018)
- 방대한 음악 데이터 셋을 결합한 음악창작 도구
- 특이한 음악 창작 도구들 개발

- Open AI

- 엘런 머스크 등이 AI의 인류에 대한 위협을 제거하고 선한 미래를 위해 투자한 비영리 연구단체  
(2015, 1억불 투자 => 2019 MS에서 1억불 추가 투자)
- Clara(2018), Musenet (2019. 5) : Transformer 구조 GPT-2 활용

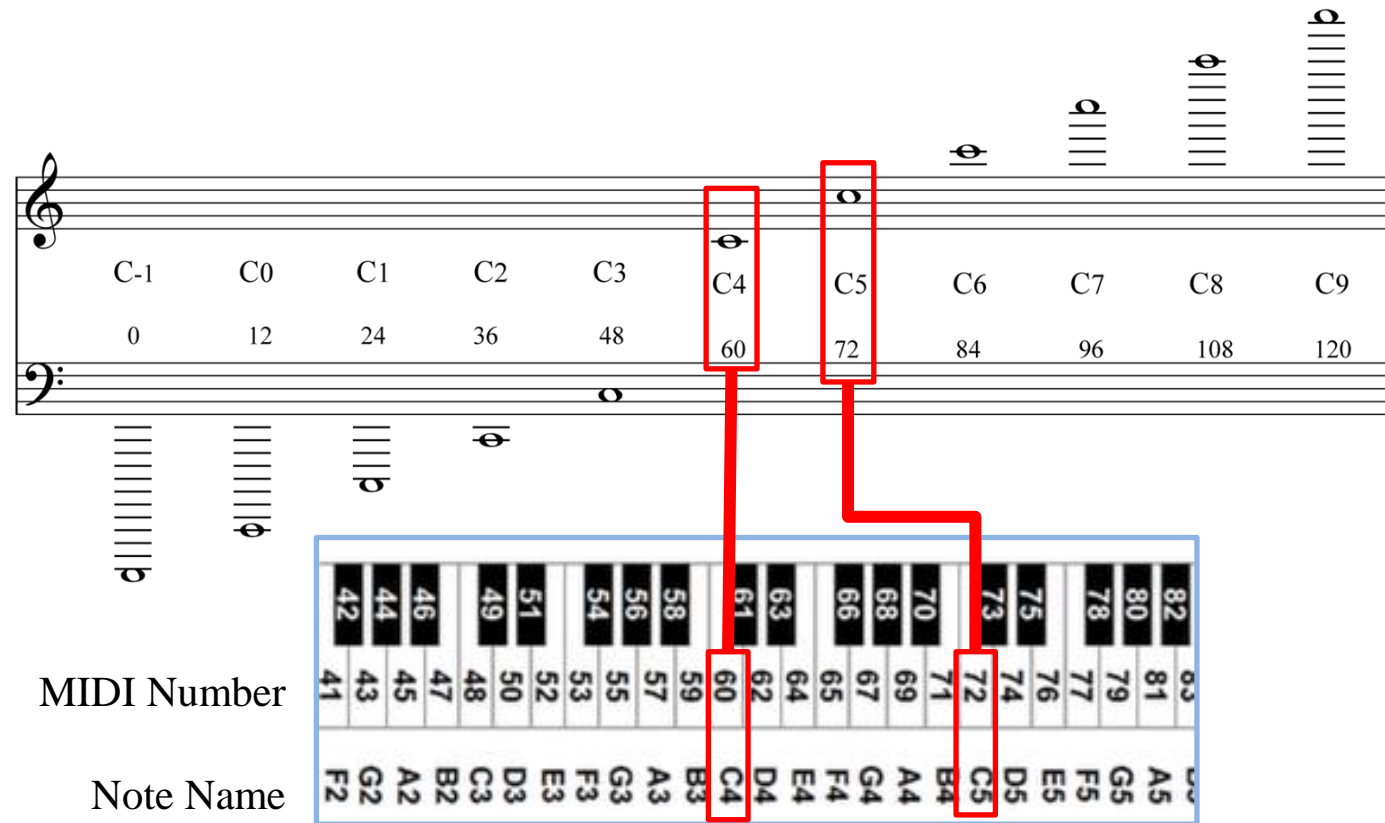
# MIDI

## Musical Instrument Digital Interface

음정, Duration, 셈여림, 등등을 Digital 화 한 protocol (since 1984)

MIDI 파일(.mid)은 기록된 정보대로 디지털 악기(음원)를 구동하여 Sound를 생성할 수 있다.(연주의개념)

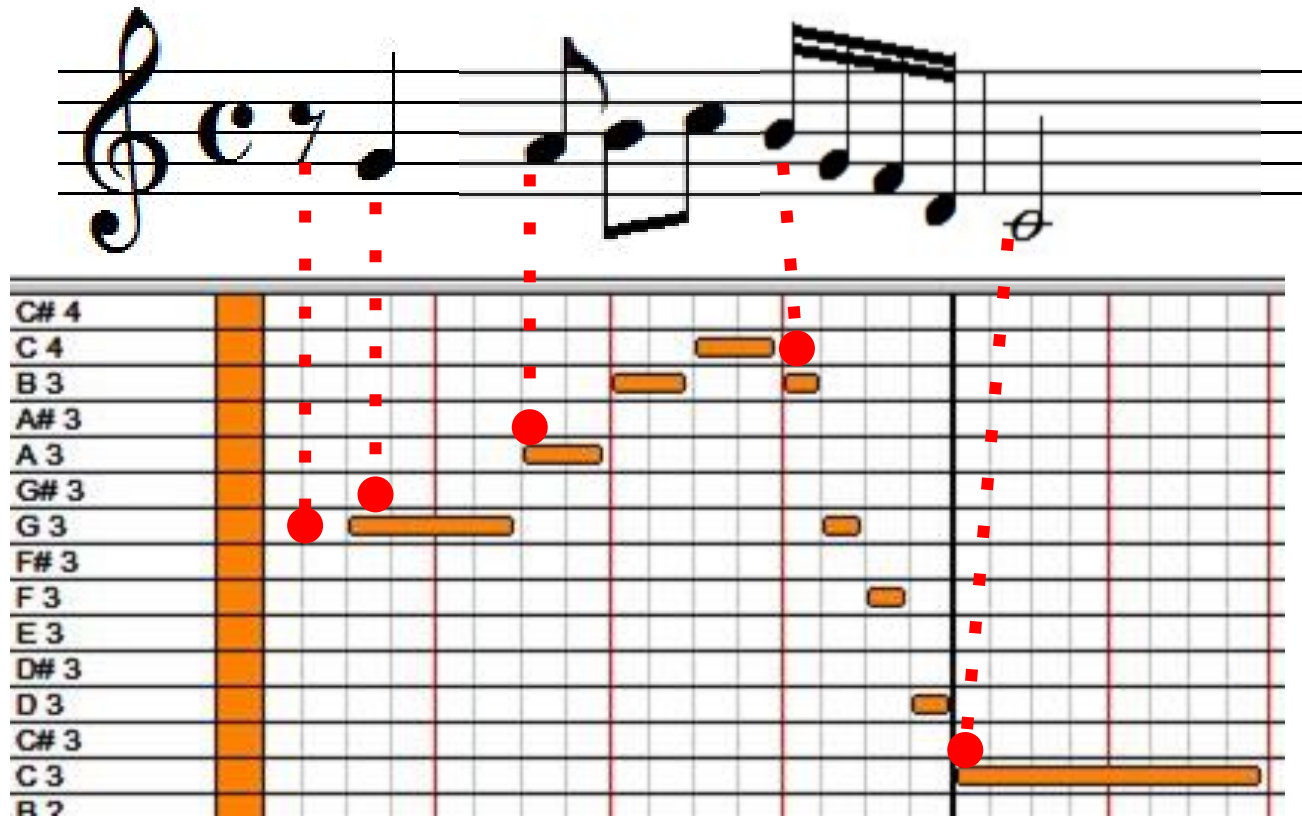
음정 : 0-127 | Duration Pulses per quarter note = max 960





# MIDI

## 변환 예시



**MIDI 변환 예시** - Time Scale 1 Bar = 192 Tick

```
<meter.TimeSignature 4/4>
```

```
<note, time24, G3, dur=48>
```

```
<note, time72, A3, dur=24>
```

...



Composing ;  
What's next?

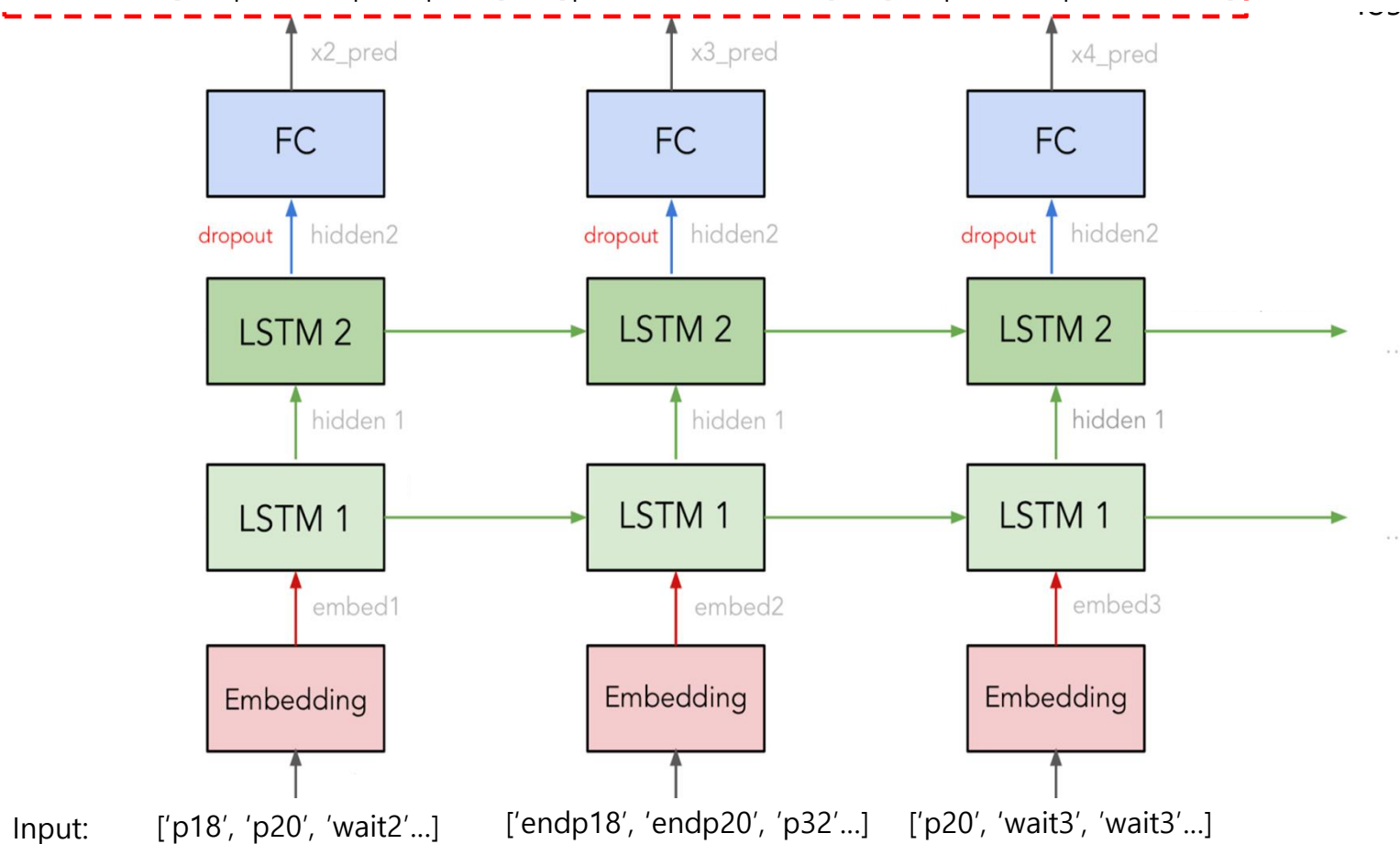
A musical score in 3/4 time, key of D major (two sharps). The score is divided into three systems. The first system (measures 1-8) is in bass clef. The second system (measures 9-10) is in treble and bass clef. The third system (measures 11-12) is also in treble and bass clef. A red rectangular box covers the right side of the third system, containing a large red question mark, indicating a point of creative decision or a placeholder for further composition.

## 4. LSTM

### Training 원리

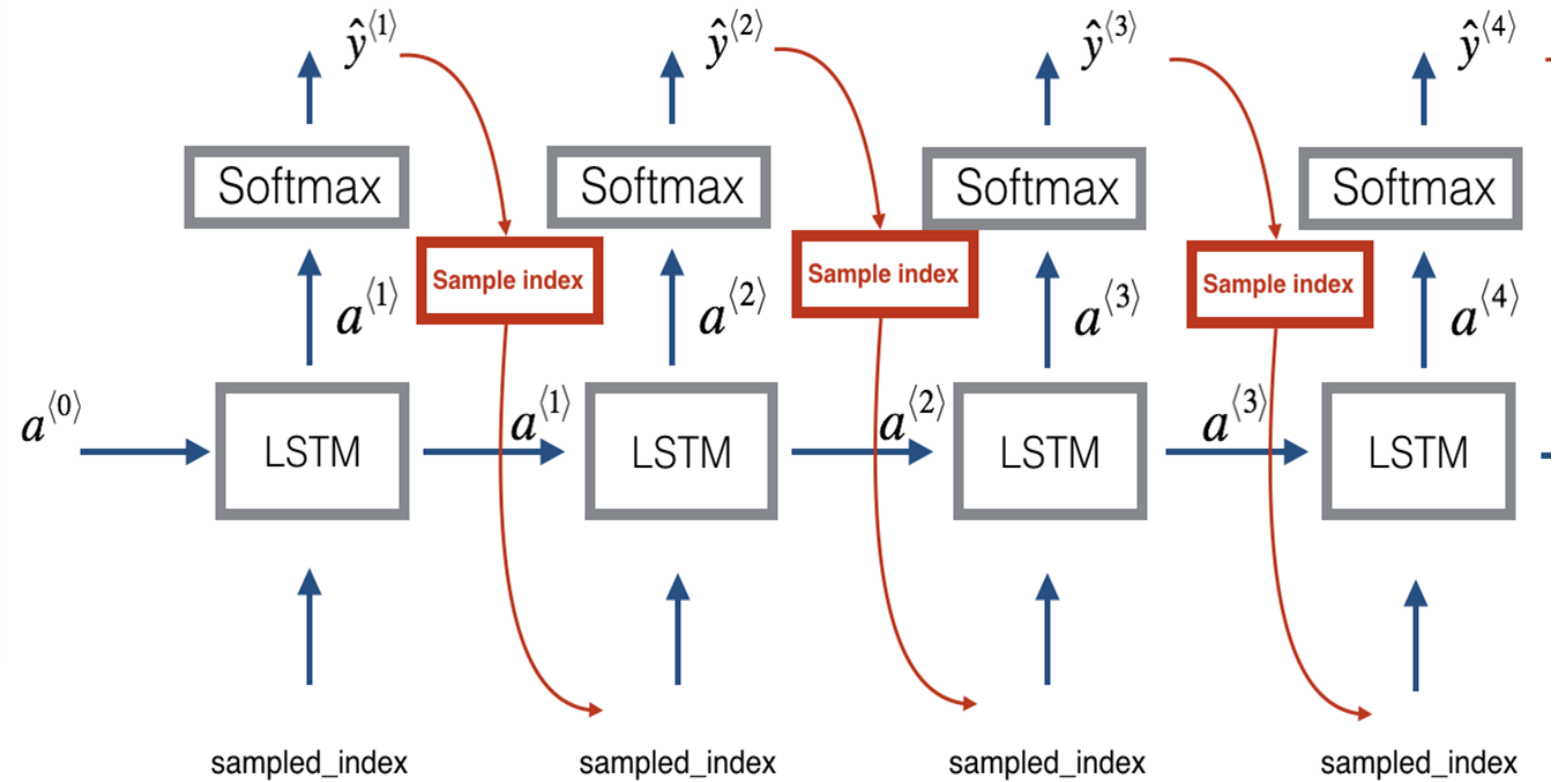
Target: ['endp18', 'endp20', 'end20'...] ['p20', 'p21', 'endp32'...] ['endp20', 'endp21', 'wait1'...]

Prediction: ['endp18', 'endp20', 'p32'...] ['p20', 'wait3', 'wait3'...] ['endp20', 'endp21', 'wait3'...]



## 4. LSTM

### 모델링 - Generating (Predict)원리



## 5. Work Flow

### ● 데이터셋 수집

- 유명 클래식 음악 작곡가 작품
- 데이터학습 가능한 MIDI 파일로 인터넷에서 수집
- <https://www.classicalarchives.com/>,  
<https://www.kaggle.com/soumikrakshit/classical-music-midi>,  
<https://composing.ai/dataset>, <https://www.classicalarchives.com/midi.html>

### ● 학습 및 모델 생성

- AI 자연어 창작 기법과 동일한 개념
- 교향곡에 최적화 된 톨 활용하여 MIDI 파일 학습, 창작
- 더 나은(그렇듯한) 결과 위한 변수 및 네트워크 튜닝 방법 연구

### ● 결과물 생성

- MIDI 화일 (음원 Wave로 변환하여 감상)

## 6. 맛보기

구글, Music Transformer 맛보기

[https://colab.research.google.com/notebooks/magenta/piano\\_transformer/piano\\_transformer.ipynb](https://colab.research.google.com/notebooks/magenta/piano_transformer/piano_transformer.ipynb)



Open AI, Musenet 맛보기

<https://openai.com/blog/musenet/>



# LSTM으로 쇼팽 느낌 피아노곡 만들기

낭만파 음악, 분위기 있는 피아노 곡  
만들기

## II-1. 구체 목표

### ● 유명 작곡가 스타일로 음악 창작하기

- 쇼팽 (Frédéric François Chopin, 1810-1849)
- 피아노의 시인으로 불리는 폴란드 작곡가
- 감성적으로 빼어난 스타일의 피아노 곡이 특징
- 쇼팽의 작품을 학습하고 새로운 스타일 곡 만들기



쇼팽 녹턴 Op.9. No.2 중에서

### ● 학습 및 모델링

- 음악 스타일 학습
- 음악적 특화된 구조와 변수 조절





## II-2. 데이터 수집 및 전처리

- Data Set - 쇼팽의 녹턴

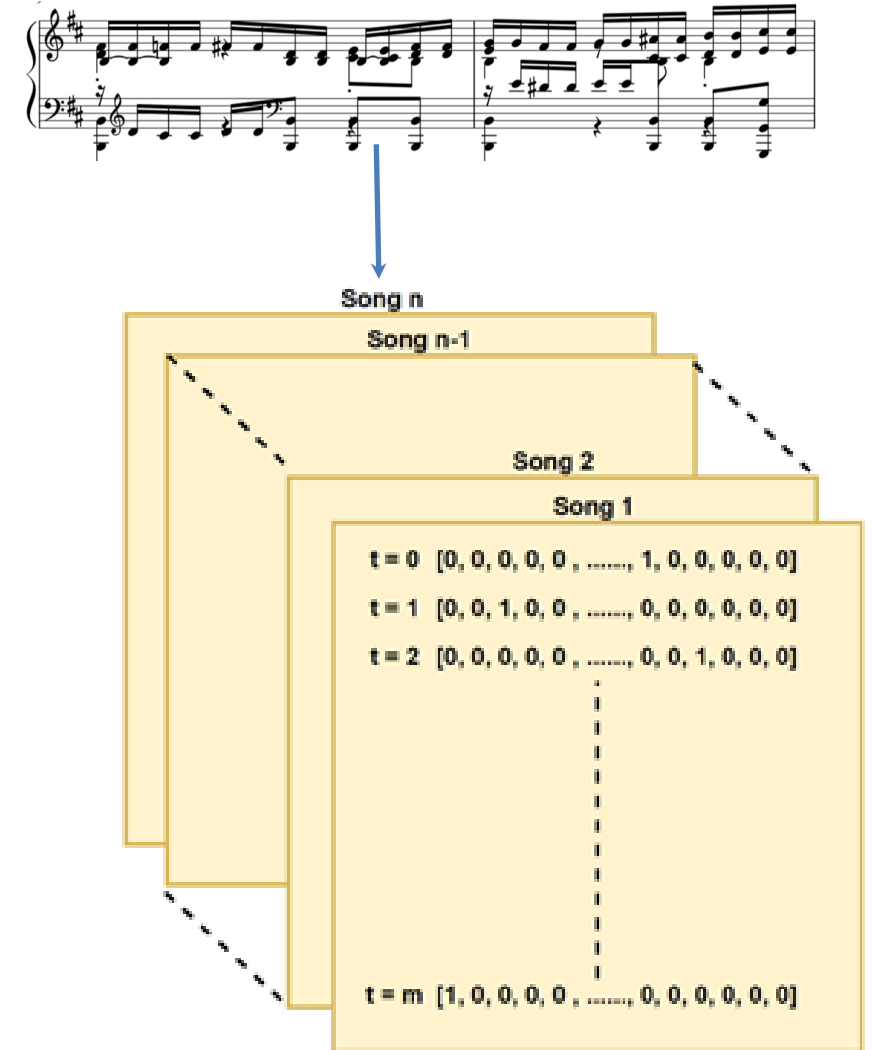
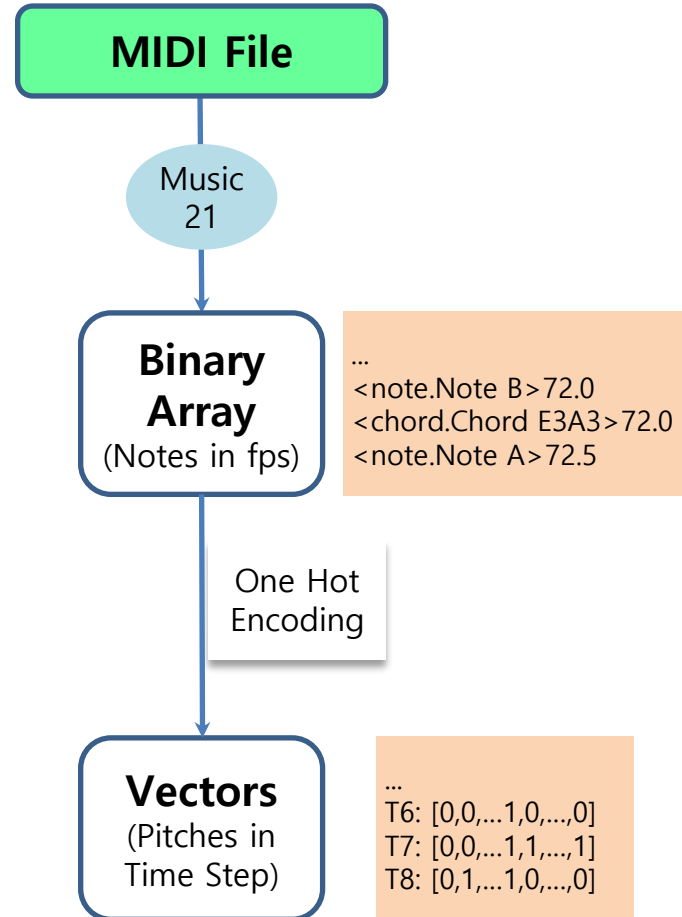
- 1827년에서 1846년까지 작곡한 21개의 곡. 독주 피아노의 연주와 연주회에서 연주되는 비중이 매우 높고 완벽한 작품으로 꼽히고 있다.



# MIDI

## - Pre Processing

### Converting MIDI Data to Vectors



- Sliding Window

seed length: 20 chracters + 1 chracter output

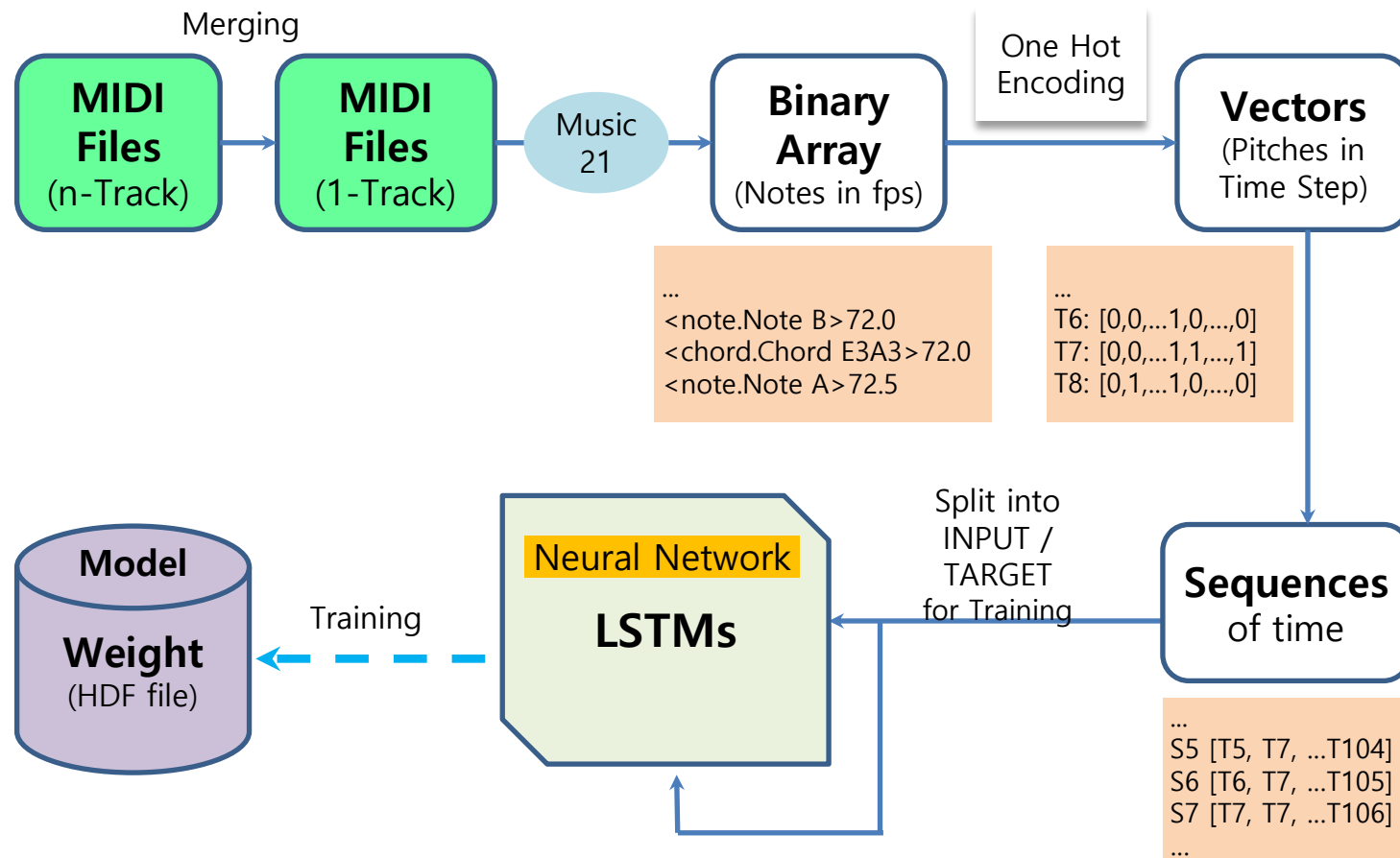
The image displays a 15x15 grid of the sentence "Tryna keep it simple is a struggle for me". Each row represents a different starting index for a sliding window of length 15. Red bounding boxes highlight the current window being processed, while green bounding boxes highlight the next window to be processed. The windows move from left to right across each row, and then from top to bottom across the grid. The sequence of windows is as follows:

- Row 1: [0, 15]
- Row 2: [1, 16]
- Row 3: [2, 17]
- Row 4: [3, 18]
- Row 5: [4, 19]
- Row 6: [5, 20]
- Row 7: [6, 21]
- Row 8: [7, 22]
- Row 9: [8, 23]
- Row 10: [9, 24]
- Row 11: [10, 25]
- Row 12: [11, 26]
- Row 13: [12, 27]
- Row 14: [13, 28]
- Row 15: [14, 29]

# II-3. LSTM

## - Training

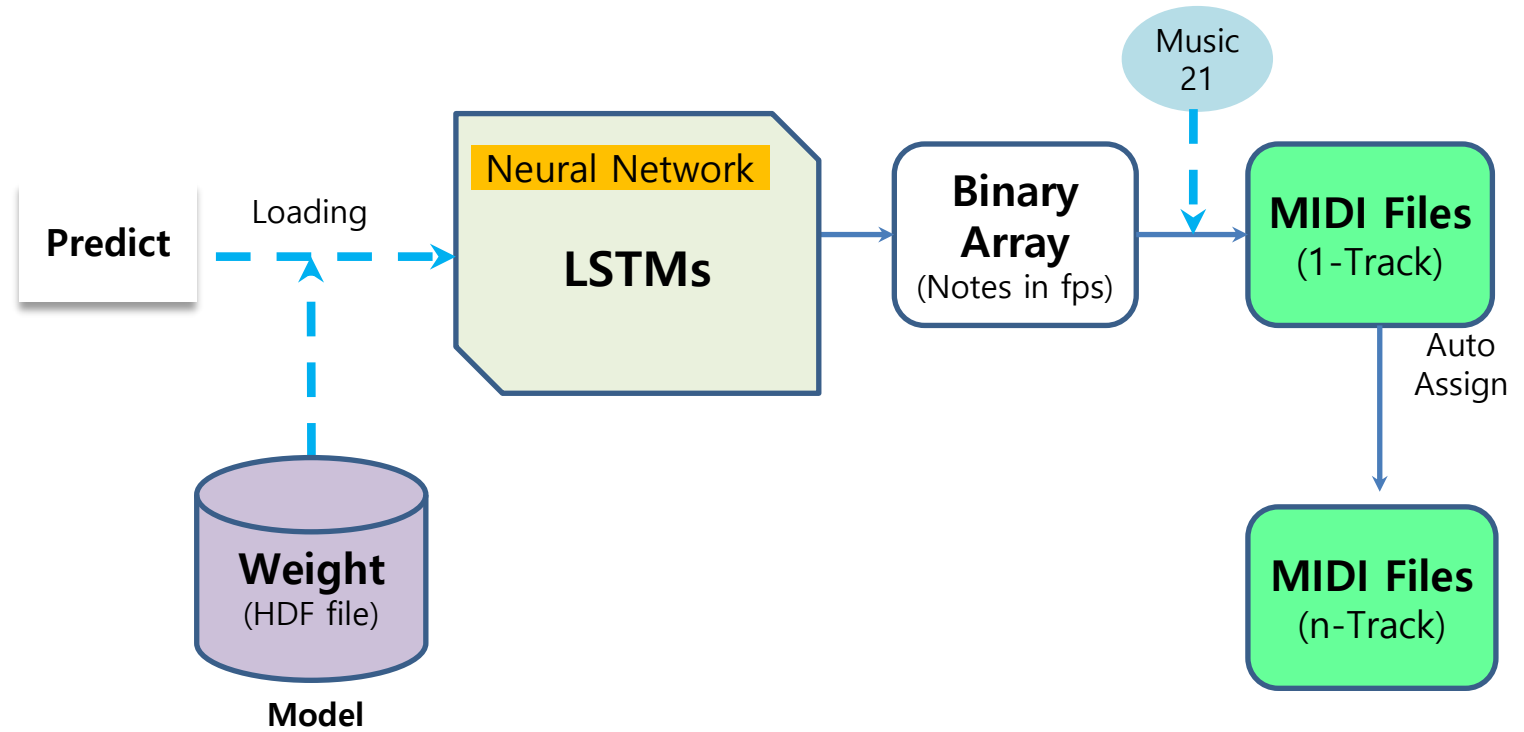
### 모델링 - Workflow



## II-4. LSTM

### - Predict

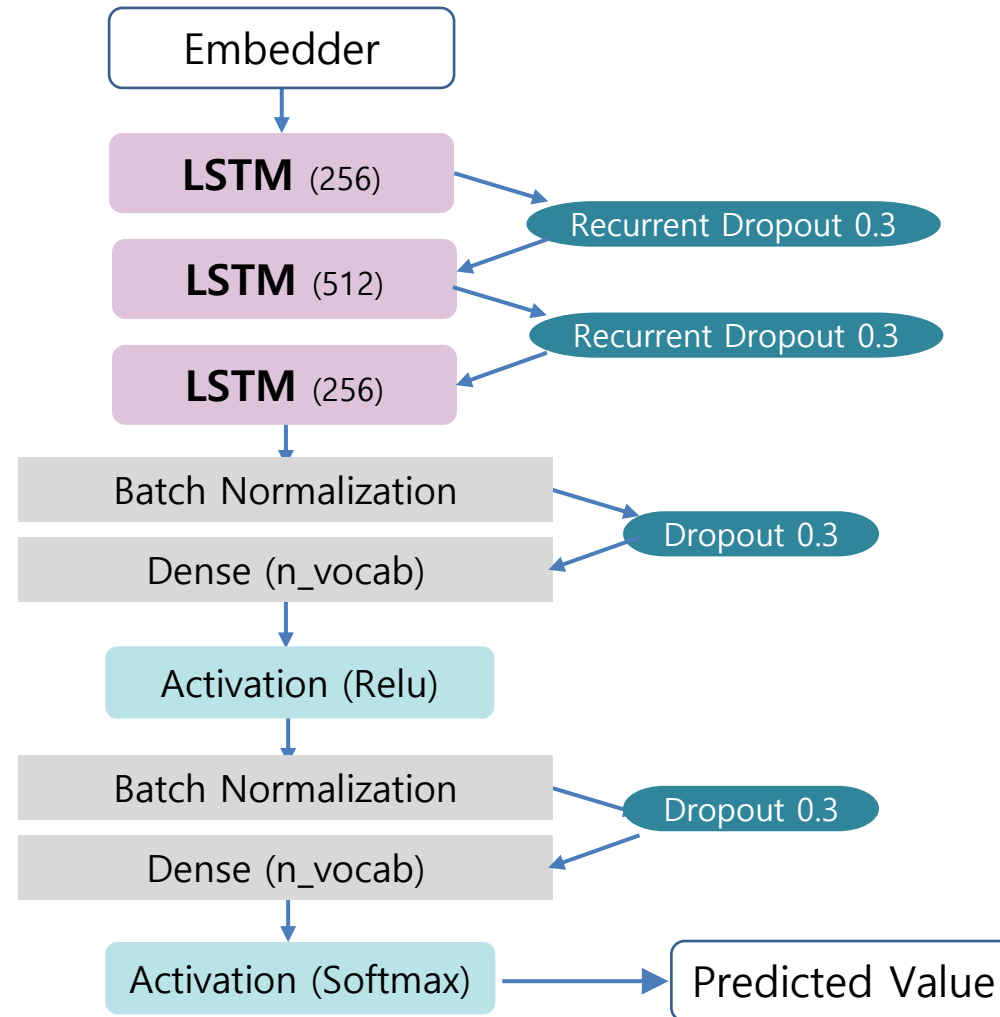
Predict - Creating & Converting to MIDI



## II-5. LSTM

### - Network

#### LSTM Network (Using KERAS)



## II-6. 실습

- Google CoLab ; Creating Music using LSTM
  - 원문 출처 (깃허브 참조)
  - 한글로 보기



# Attention 활용한 교향곡 창작 예시

복잡하고 긴 클래식 음악의 최고봉,  
교향곡

# III-1. 구체 목표

## ● 유명 작곡가 스타일로 음악 창작하기

- 슈베르트, 미완성 교향곡
- Symphony No.8 'Unfinished' D.759
- 작곡가 죽기 6년 전인 1822년 작곡
- 1, 2악장만 있고 3, 4악장이 없음(원래 4악장)
- 슈베르트 작품 등등을 학습하여 3악장 완성



미완성교향곡 1악장 중에서



## ● 학습 및 모델링

- 악기 편성과 편곡 스타일 학습
- 음악 장르, 특성(ex 소나타 형식, 낭만파)에 특화된 구조와 변수 조절

## III-2. 데이터 수집 및 전처리

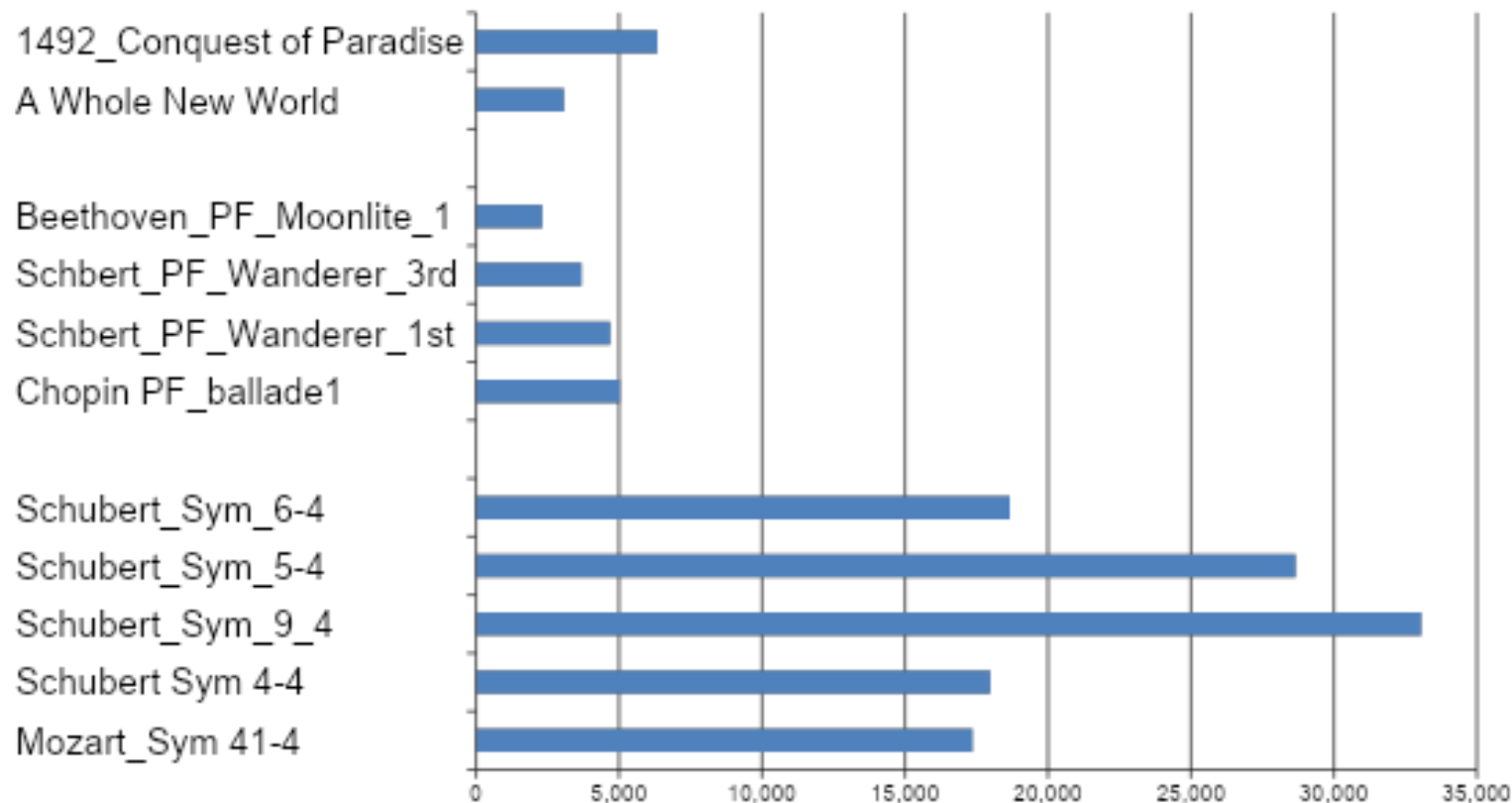
### ● Data Set

장르	작곡가	곡명	곡수	비고
교향곡	슈베르트	No.1 ~ 9	8	
관현악곡		로자문데	5	
실내악곡		현악 4중주	10	
		피아노 소나타	12	
		기타 실내악곡	32	
소계			67	
계			684	12음계 전조

교향곡	하이든	No.81 ~ 104	24	전성기
	모차르트	No.25 ~ 41	17	전성기
	베토벤	No.1 ~ 8	8	9번 제외
관현악곡	베토벤	피델리오 등	5	슈베르트에 큰 영향 준 곡
계			54	

## III-2. 데이터 수집 및 전처리

Music Notes in a Song



# III-2. 데이터 수집 및 전처리



## Symphonie in H moll (unvollendet) von FRANZ SCHUBERT

Allegro moderato.

Flauti.  
Oboi.  
Clarinetti in A.  
Fagotti.  
Corni in D.  
Trombe in E.  
Alto, Tenore.  
Tromboni.  
Basso.  
Timpani in H.Fis.  
Violino I.  
Violino II.  
Viola.  
Violoncello.  
Basso.

1

연주 쉬는  
악기들

12

2

18

3



## III-2. 데이터 수집 및 전처리

### ● 정확한 학습과 모델링을 위해 데이터 축약

14개 악기 악보 => 피아노 용 연주곡 형태로 Merge

Allegro moderato

Flauti.

Oboi.

Clarinetti in A.

Fagotti.

Corni in D.

Trombe in E.

Alto.  
Tenore.

Tromboni.

Basso.

Timpani in H.Fis.

Violino I.

Violino II.

Viola.

Violoncello.

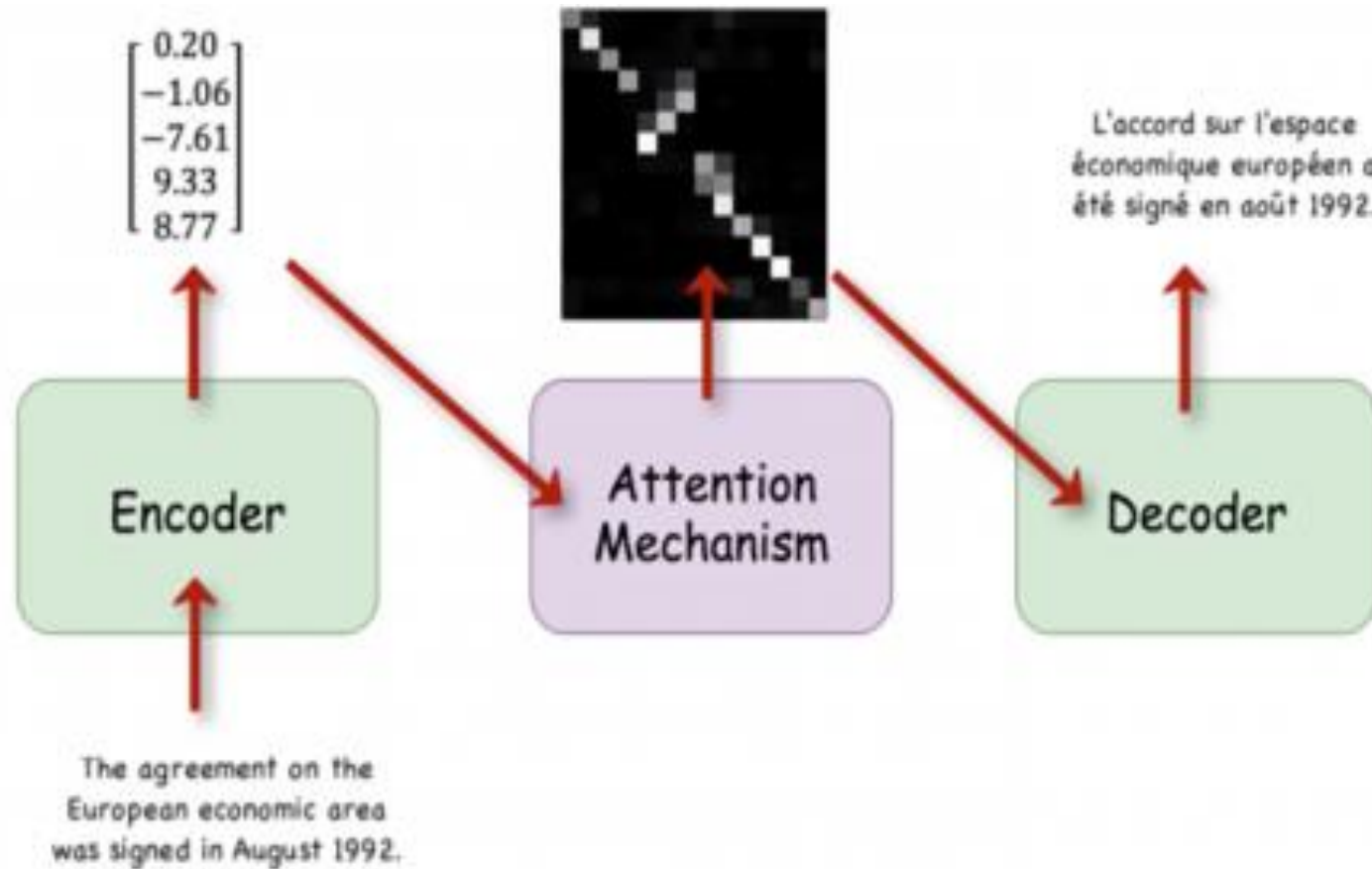
Basso.

9

11

## III-4. Self Attention

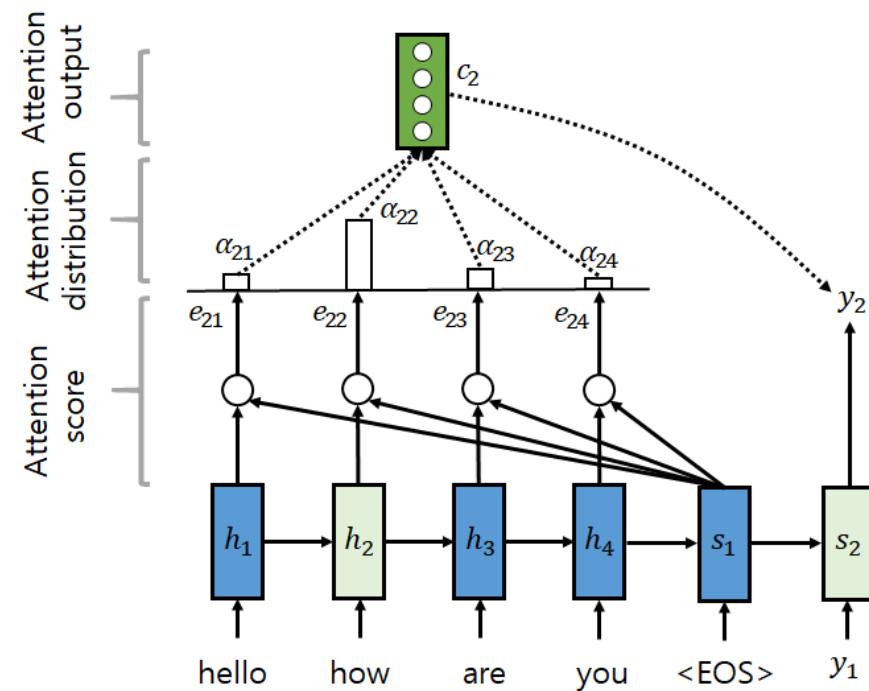
Seq2Seq에서 Context Vector 대신 Attention 기법을 쓰는 것



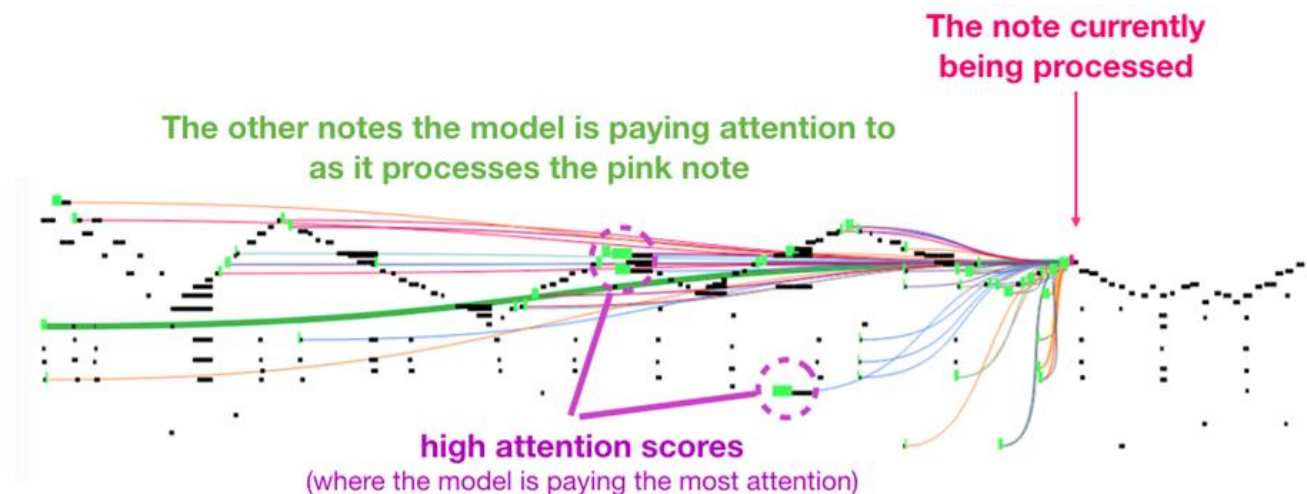


# III-4. Self Attention

Attention Concept

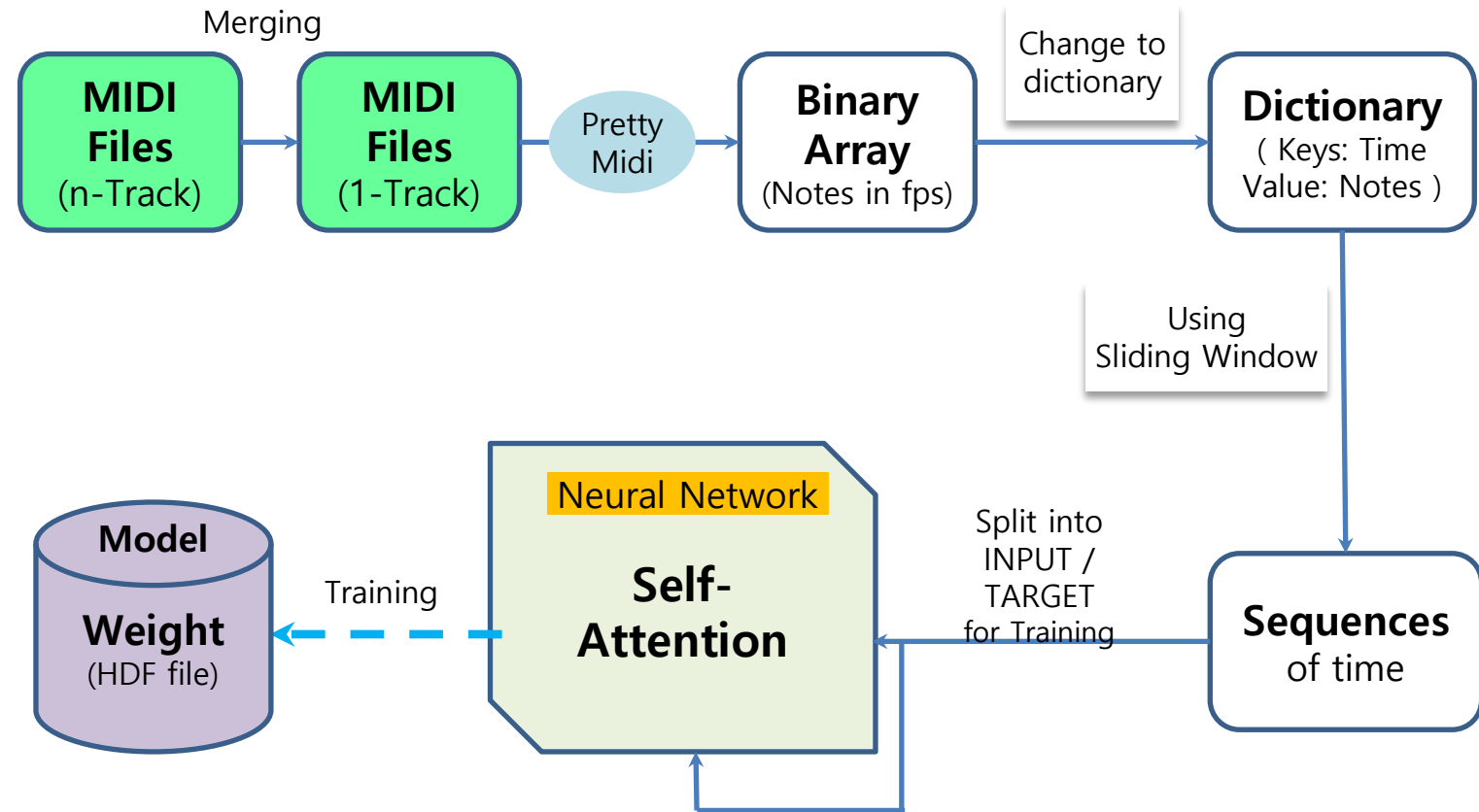


Self Attention



# III-5. Self Attention -Training

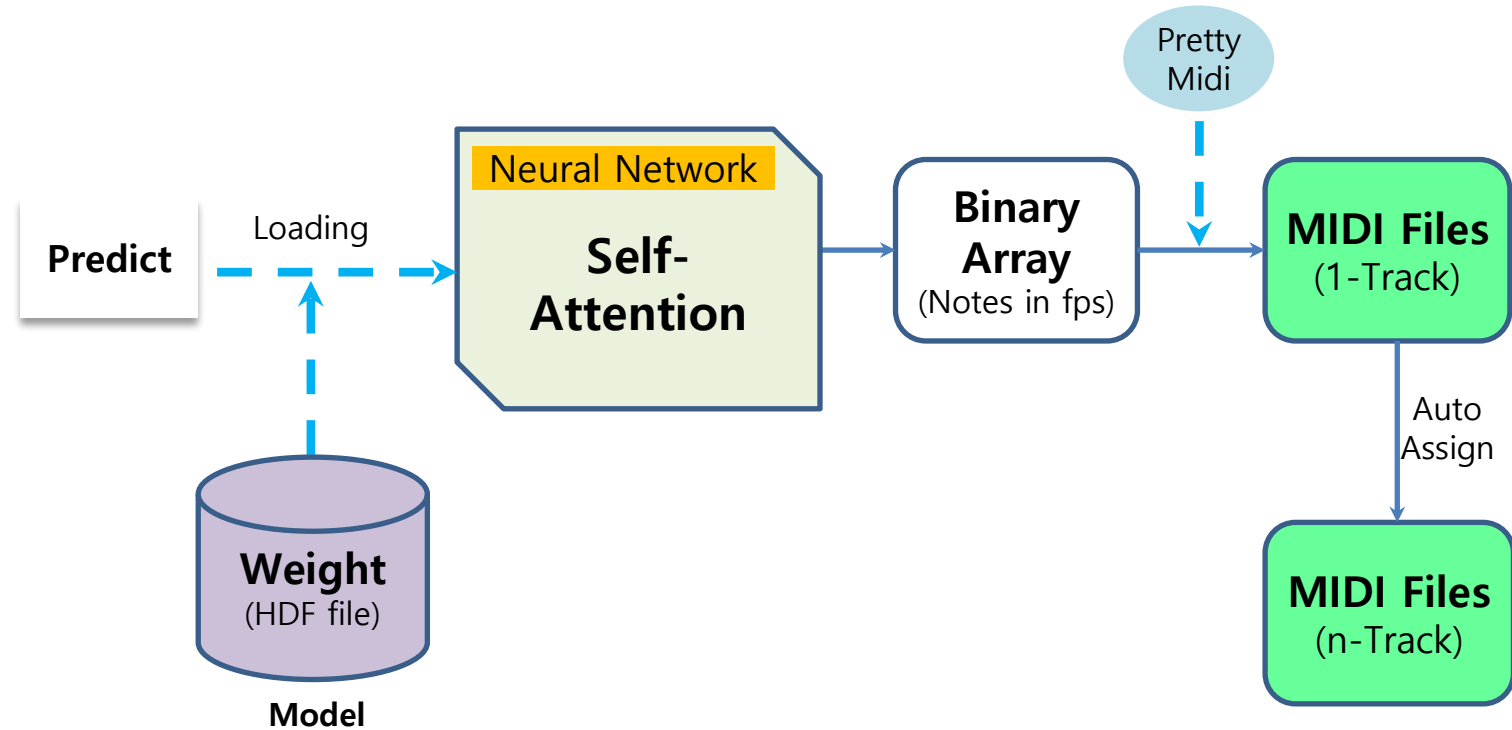
## Training Work Flow



## III-6. Self Attention

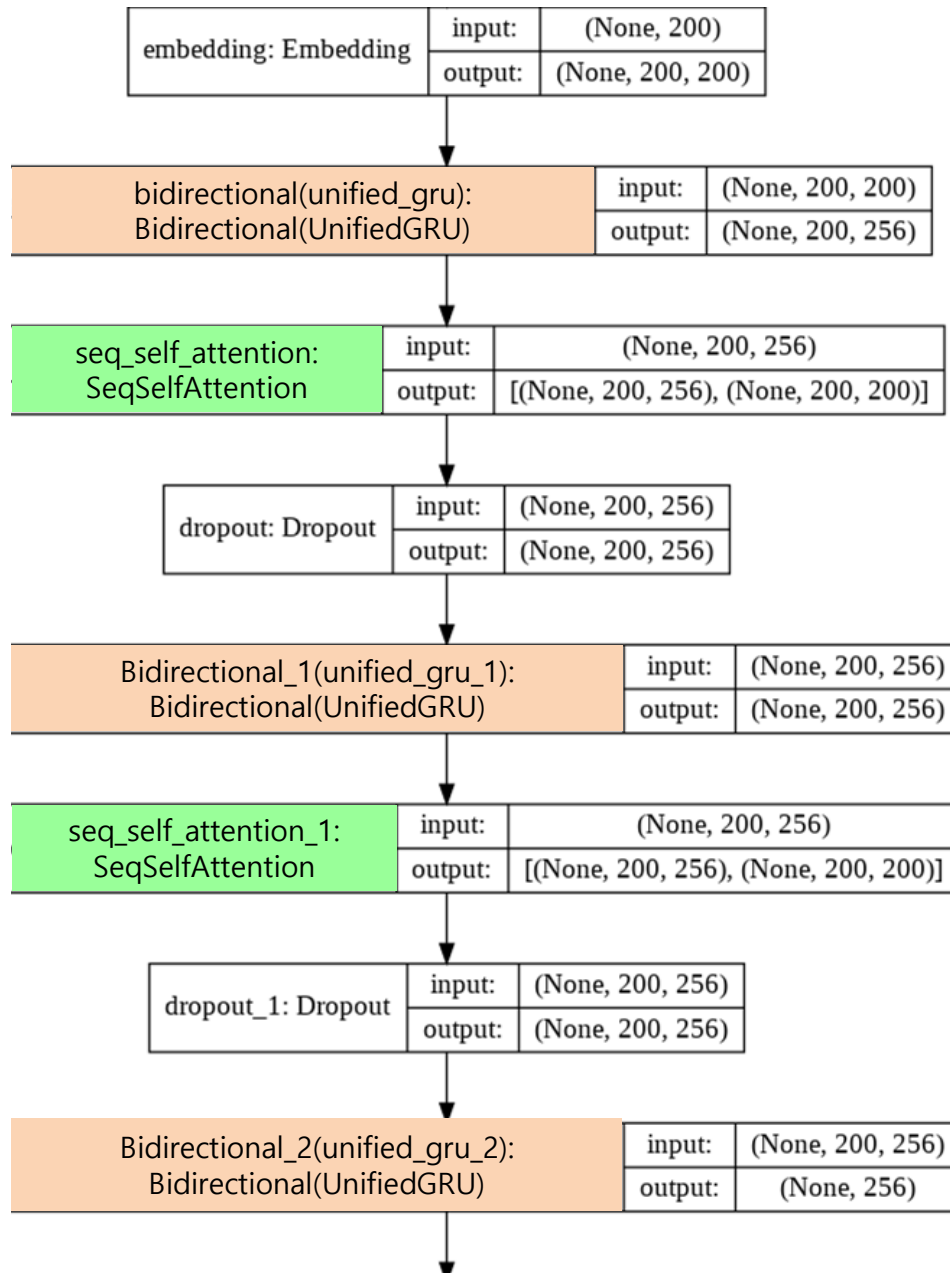
### - Predict

#### Generating WorkFlow

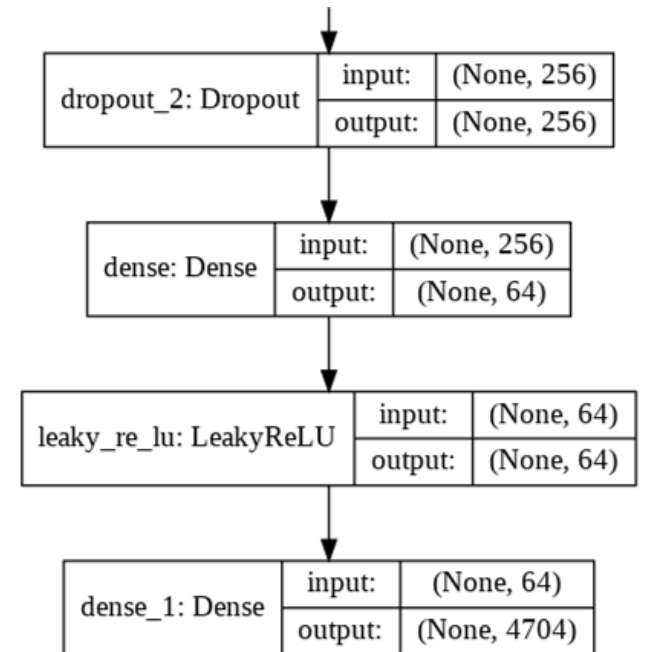


# III-7. Self Attention

## - Network



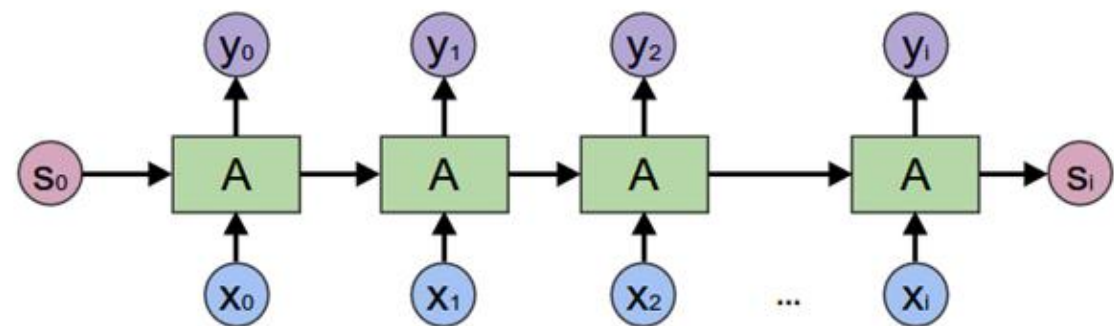
## Network Using Tensorflow2.0



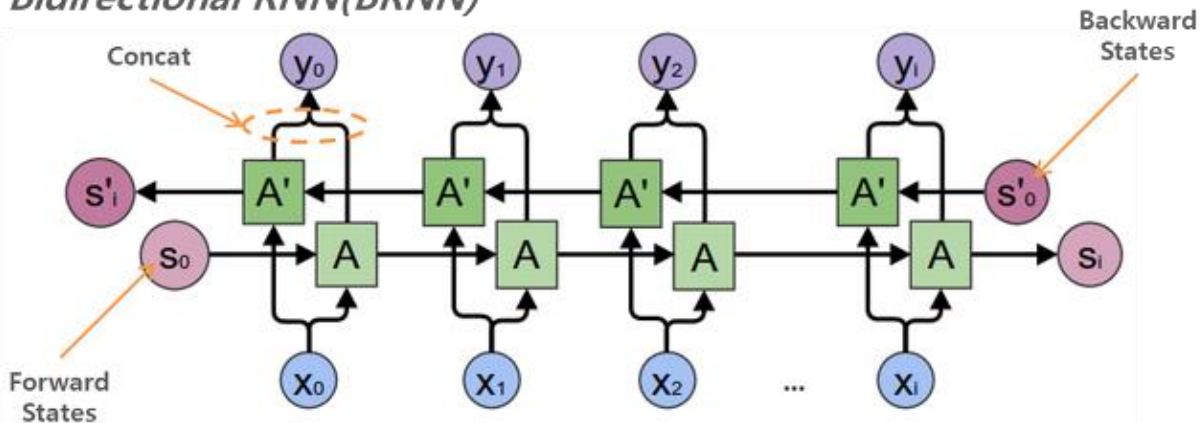
- GRU는 복잡한 LSTM의 구조를 간단히 한 것으로 LSTM과 유사한 성능을 가집니다.

# BiDirectional RNN

*Basic RNN*



*Bidirectional RNN(BRNN)*



- BRNN은 양방향으로부터 시퀀스의 **과거 정보**( $t$  기준  $t-1, t-2, \dots$ )와 **미래 정보**( $t$  기준  $t+1, t+2, \dots$ ) 둘 다 사용하기 때문에 기존 RNN 보다 더 좋은 성능을 기대할 수 있습니다.

## III-8. 실습

- Google CoLab ;
  - GRU + Self Attention Method for Creating Music



# 결과 ; 음악 창작에 어울리는 도구 분석

Attention사용하는 Transformer

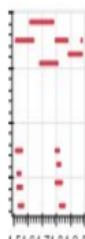


## IV. 기법별 비교

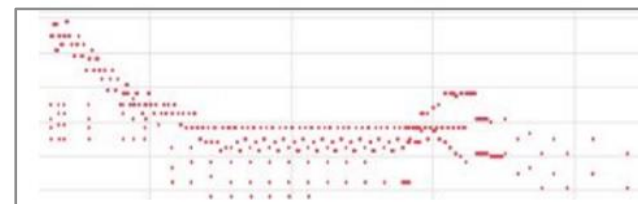
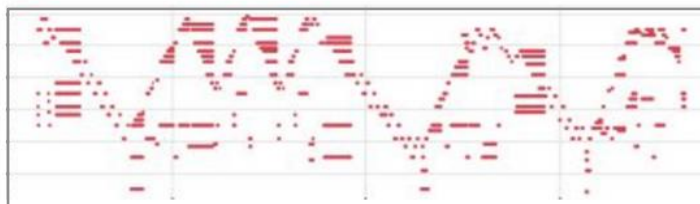
### ● 시간흐름에 따른 알고리즘 별 생성물 비교 (by Google Magenta)

- Attention 기법의 Music Transformer가 긴 시간 동안 유효성 유지

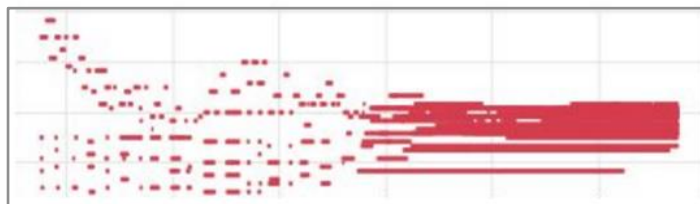
Prime



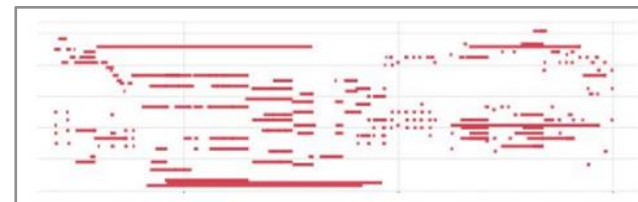
Music  
Transformer



Baseline  
Transformer



LSTM



Epoch=500, Loss = 0.4 (24hours in CoLab)

# IV-1.결과

## - LSTM



Epoch = 300, Loss= 0.5 (24hours in CoLab)

## IV-2.결과

### -GRU & ATTENTION



A musical score for piano, consisting of four systems of staves. The first system is marked with a tempo of ♩ = 120. The score is written in treble and bass clefs with a key signature of two sharps (F# and C#). The notation includes various musical symbols such as notes, rests, and dynamic markings. The second system starts with a measure number of 7, the third with 11, and the fourth with 14.

## IV-3. 결과 평가

### 결과물 평가

- **Numerical Evaluation**

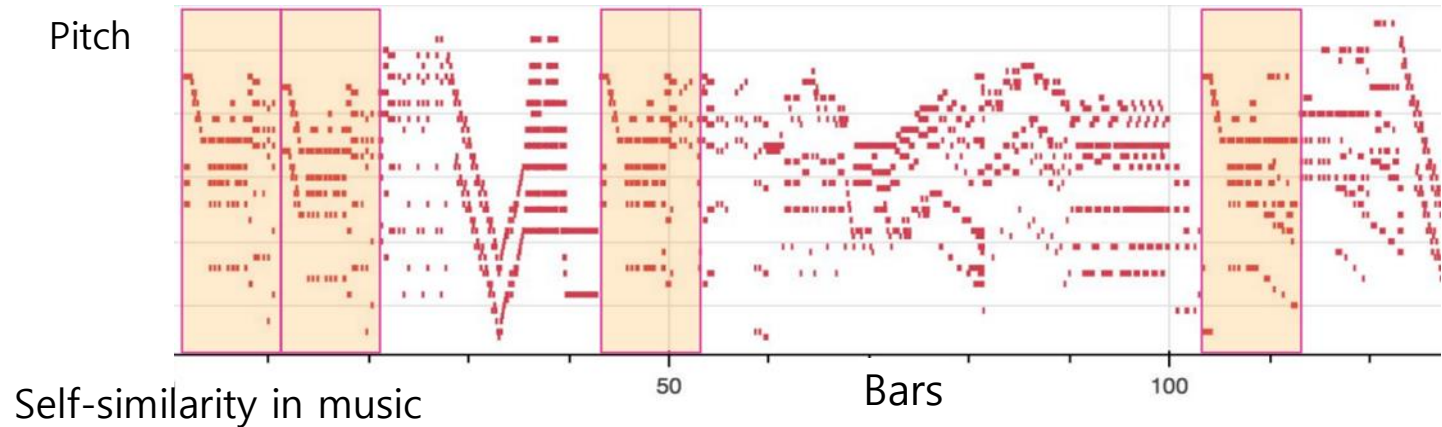
- Empirical Entropy 측정
- Cross Entropy 측정
- 여러 악기가 사용되는 교향곡의 경우 적합하지 않음

- **Listening Evaluation**

- 생성된 MIDI 파일을 음원으로 변환
- 피아노 트랙을 간단한 현악기군으로 변환
- Turing Test/ Blind Test 실시 음악적 완성도 평가
- 작곡가, 평론가, 지휘자, 연주자 16명 대상

## IV-3. 결과 평가

- LSTM은 많은 연산시간과 Long Term Dependency에 문제
- GRU + Self ATTN 방식이 생성 길이가 길어짐에도 음악적 구조와 문법을 자연스럽게 유지함



Motifs repeat, immediately and also at a distance

- Self ATTN 방식은 Transformer / BERT에도 차용되어 음악 창작, 문장 창작 및 번역에 기여