

RNN Auto-Encoder의 시계열 임베딩을 이용한 자동 작곡

최수지

INDEX

1. Introduction
2. RNN Auto-Encoder
3. Experiment
4. Conclusion

1. Introduction

음악 작곡

기존 연구 : 전방향 인공신경망(Feedforward Artificial Neural Network)

- 시계열 데이터 학습 어려움
- 학습할 곡의 데이터를 회귀적으로 구성
- 학습 데이터가 많은 경우 각각의 곡이 적절히 임베딩 되지 않음

RNN Auto Encoder 신경망 구조

- RNN + AutoEncoder
- 시계열 학습 가능
- 입력데이터 특성 추출로 다양한 특성을 가진 곡 작곡

2. RNN Auto-Encoder

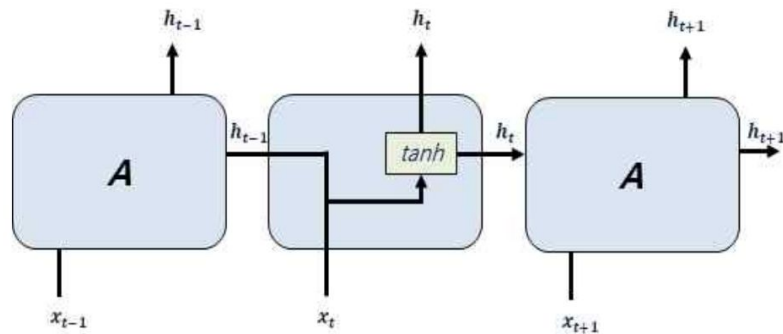
RNN

- 은닉층이 순환하는 구조를 가진 신경망
- 은닉층의 상태가 현재입력과 이전 은닉층의 상태로 결정
- 단기적인 의존관계만 학습 → 장기적인 의존관계 학습 불가

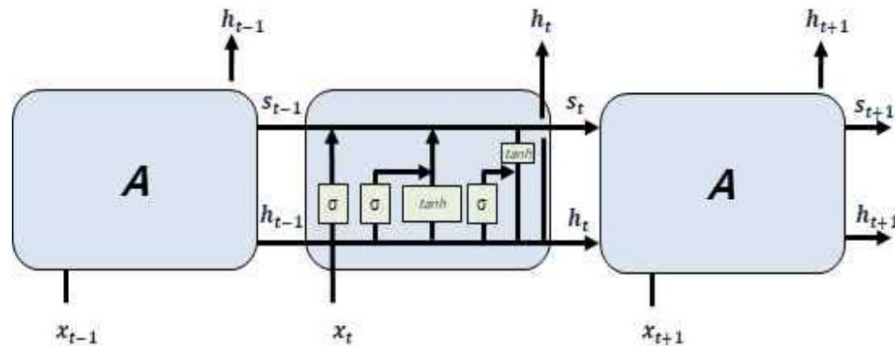
LSTM

- RNN의 종류
- 장기의존성 문제를 해결

2. RNN Auto-Encoder



일반RNN

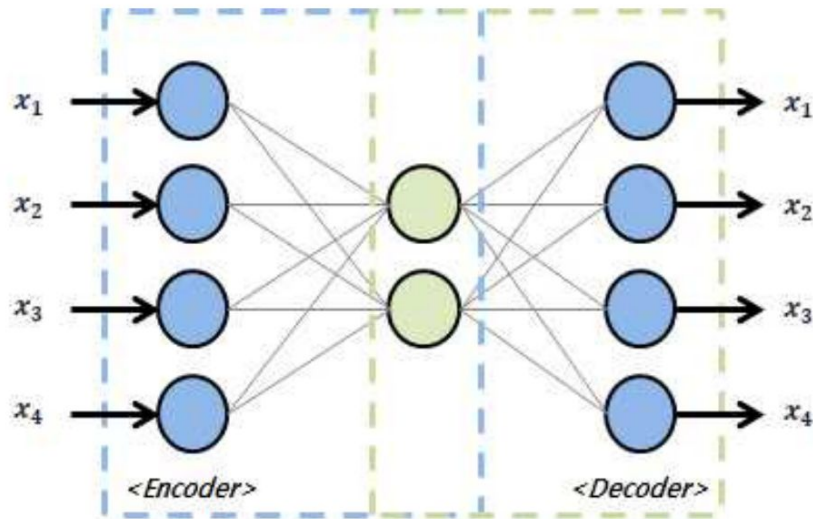


LSTM

2. RNN Auto-Encoder

Auto-Encoder

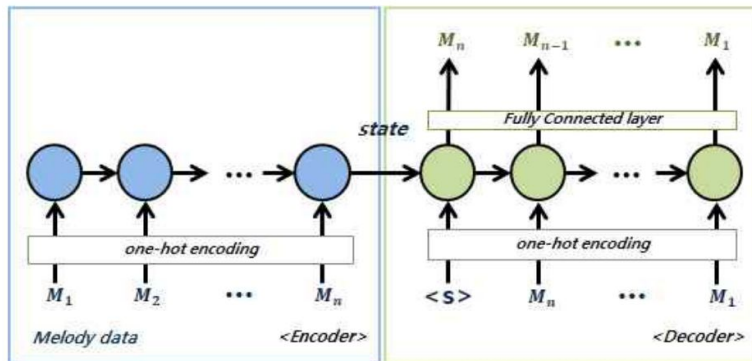
- 입력데이터로만 학습하여 은닉층에 데이터의 특징을 압축시키는 비지도 학습
- Encoder와 Decoder로 구성
- 대칭적인 구조
- 높은 차원의 입력데이터를 낮은 차원으로 변환
->입력 데이터의 특징 추출



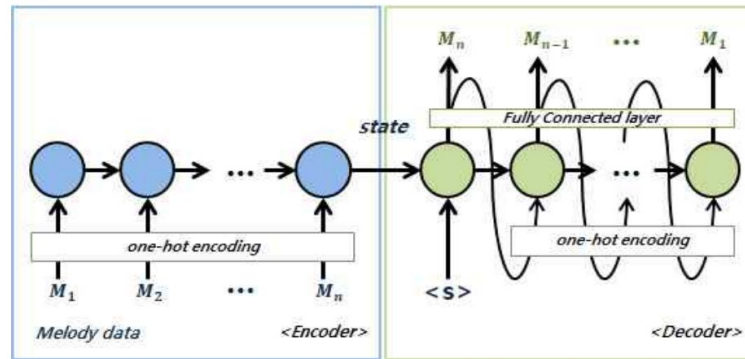
2. RNN Auto-Encoder

RNN Auto-Encoder

- RNN과 Auto-Encoder를 결합한 새로운 구조의 신경망
- RNN Encoder는 one-hot encoding 된 시계열데이터가 입력되어 학습
- RNN Decoder는 시작태그와 함께 입력 시계열데이터의 역순으로 출력되게 학습



<학습>



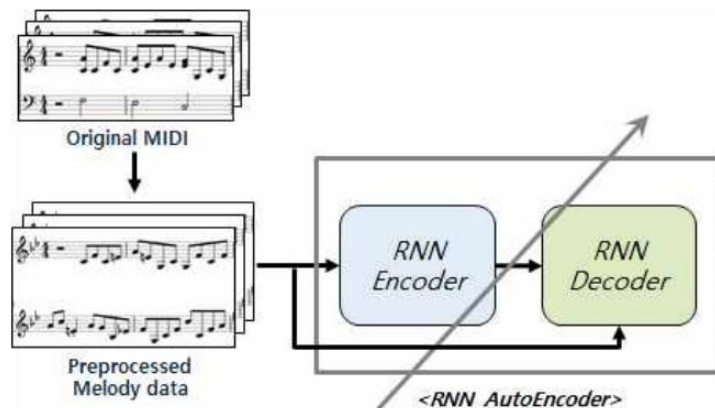
<작곡>

3. Experiments

RNN Auto-Encoder를 이용한 작곡

<데이터 학습>

1. midi 파일 전처리 : 화음에서 근음 추출, 음표/쉼표/박자 추출
2. One-hot encoding 된 시계열데이터를 RNN Encoder의 입력으로 하고 동일한 데이터를 RNN Decoder의 출력으로 하여 학습
3. 같은 방법으로 다수의 곡을 학습할 수 있으며 학습이 종료된 후에 작곡 가능

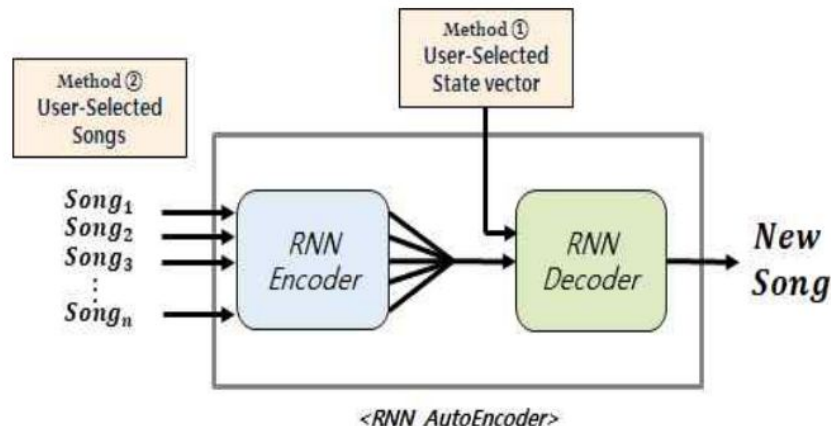


3. Experiments

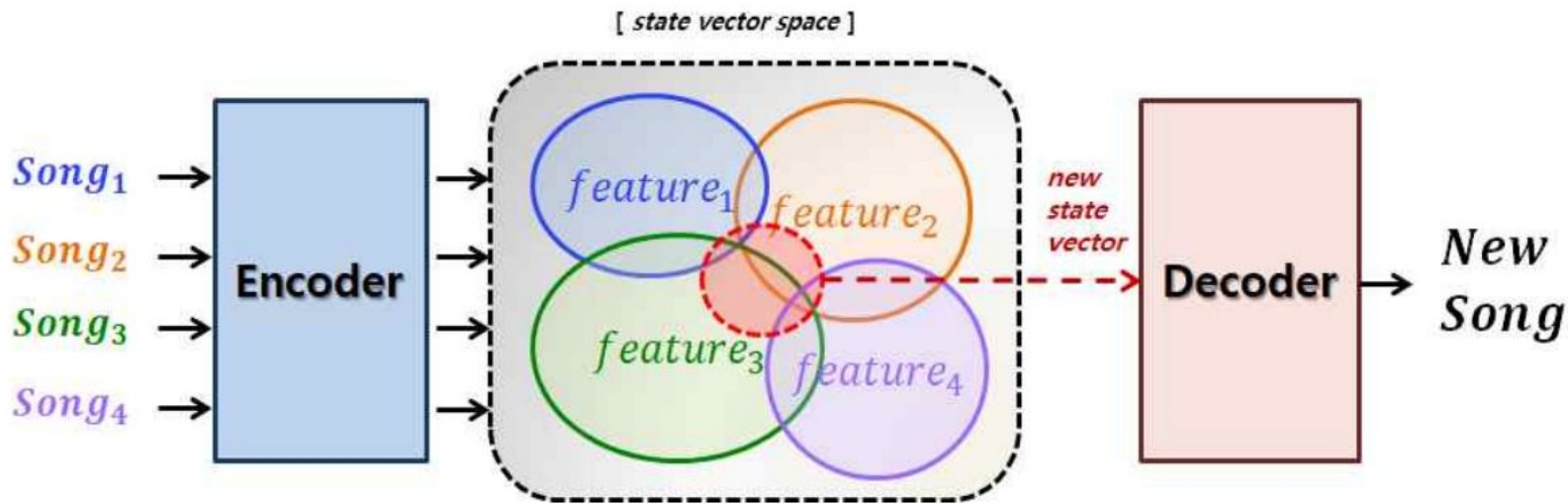
RNN Auto-Encoder를 이용한 작곡

<작곡>

1. 시작태그와 사용자가 선택한 상태벡터를 RNN Decoder에 입력
 - 사용자가 임의로 작곡용 상태벡터 설정
 - 기존 곡의 상태 벡터를 이용
2. 학습한 곡을 여러 개 선택하고
해당 곡의 RNN Encoder 출력 상태 이용
 - n개의 곡을 선택한 경우 :
출력상태의 평균을 작곡용 상태벡터로 사용
-> 선택한 곡들의 특징이 작곡된 곡에 반영



2. RNN Auto-Encoder



- 4개의 곡의 평균 상태 벡터로 작곡 할 경우
- 곡을 대량으로 학습할수록 상태 공간에 매핑되는 곡 증가
-> 더 다양한 특징이 반영된 음악 작곡

3. Experiment

RNN Auto-Encoder 작곡

- 텐서플로우 상에서 구현
- RNN Auto-Encoder의 RNN Encoder와 RNN Decoder는 128개의 은닉노드를 갖는 구조
- 하나의 곡을 학습한 후 해당 곡의 상태벡터를 넣어주면 원곡이 그대로 출력
- 다수의 곡을 학습한 후 학습한 곡 중 두 곡의 평균상태벡터를 넣어주면 두 곡의 특징이 적절히 섞인 곡이 출력

3. Experiment



(a)



(b)



(c)

- (a)와 (b)는 학습한 곡
- (c)는 (a) 와 (b) 두 곡의 상태벡터의 평균으로 작곡한 결과
- 작곡된 곡이 학습된 두 곡의 특징을 모두 반영

3. Experiment



26개의 곡을 학습하고 임의의 상태벡터를 넣어 작곡한 결과

4. Conclusion

- RNN Auto-Encoder를 이용한 자동작곡 방법 제안
- RNN Auto-Encoder를 구현하여 곡을 학습
- 곡을 학습하면 동일한 곡을 출력하는 상태벡터로 해당곡이 임베딩
- 다수의 곡을 학습한 후 그 중 두 곡의 상태벡터 평균을 RNN Decoder에 입력하여 두 곡의 특징이 나타나는 곡 출력
- 다수의 곡을 학습한 후에 임의의 상태벡터를 RNN Decoder에 입력하여 기존 곡들의 상태벡터와 작곡을 위해 넣어준 상태벡터의 특징을 모두 갖는 새로운 곡을 출력



THANK YOU