

운동주체 시 생성기

운동주체 시 생성기는 시작 단어와 총 단어수를 정해주면 운동주 시인의 문체로 시를 써주는 프로그램이다.

Features

운동주체 시 생성기는 딥러닝 LSTM(Long Short Term Memory) 모델을 통해 운동주 시인의 문체를 구현했다. LSTM 모델은 과거의 데이터가 미래에 영향을 줄 수 있도록 설계된 RNN(순환신경망, Recurrent Neural Network)의 발전된 형태이다. 본 프로젝트에서는 운동주 시인의 시 세계 속 문맥을 반영하여 텍스트를 생성하는 LSTM 모델을 구현하였다. 모델의 학습 데이터는 운동주 시인의 유고 시집인 '하늘과 바람과 별과 시'로 1941년에 출간될 예정이었던 19편의 시들이다. 운동주 시인은 사망한지 50년이 지났기 때문에 시에 대한 저작권산권은 소멸되었다. 따라서 동일성을 유지한 상태로 데이터를 활용하는 경우에는 법적인 문제가 발생되지 않는다. poem.csv 파일에 저장된 시의 제목 및 본문 데이터는 총 223행의 문장들이다. LSTM 모델이 운동주 시인의 독특한 문체를 학습하기 위해서는, 전체 문장 집합에 있는 단어들을 모두 고려하여 데이터를 재구성해야 한다. 이를 위해 재구성된 단어 집합의 크기는 총 794단어다. 데이터에 대한 원핫인코딩을 수행하는데, 가장 긴 문장의 길이는 76이고 0을 사용하여 모든 문장들의 길이를 가장 긴 문장의 길이로 맞춰주었다. 각 단어의 임베딩 벡터가 10차원이고 128의 은닉 상태 크기를 가지는 LSTM 모델을 생성하였다. 그리고 데이터를 학습시켜서 예측 모델을 만들고, 새로운 문장을 생성하는 함수를 정의해두었다. 사용자가 시를 시작할 단어와 시의 총 단어수를 입력하면, 이에 맞도록 운동주 시인의 문체로 시를 생성하여 결과를 출력해준다.

Requirements

- Python + Pandas + Numpy
- Tensorflow + Keras

Installation

- [Anaconda](#) 설치
- Anaconda Prompt 실행
- Anaconda 업그레이드, Tensorflow 설치, Keras 설치 등을 위한 아래 명령어들을 실행

```
conda update -n base conda
conda update --all
python -m pip install --upgrade pip
conda install tensorflow
conda install keras
```

How to use?

- 아래 명령어로 Python 코드 실행

```
python poem_generator.py
```

- Epoch 200/200 이 될 때까지 학습 대기
- 어떤 단어로 시를 시작할 것인지 단어 입력
- 몇 단어로 시를 쓸 것인지 숫자 입력
- 운동주체 시 생성 결과 확인

Sample Output

운동주체 시 생성기가 어떤 단어로 시를 시작할 것인지 단어를 입력해주세요.

사랑

운동주체 시 생성기가 몇 단어로 시를 쓸 것인지 숫자를 입력해주세요.

24

'사랑' 단어로 시작하는 총 24 단어의 운동주체 시 생성 예제

사랑 돌과 돌과 돌이 끝없이 연달아 언덕 못할 마음으로 금잔화 한 포기를 따 가슴에 꽂고 병실 안으로 사라진다 나는 그 여자의 건강이 아니 내 건강도

'나는' 단어로 시작하는 총 24 단어의 운동주체 시 생성 예제

나는 이마에 땀을 흘려야겠다 여미고 화단에서 금잔화 내려 슬픈 것처럼 안에까지 눈이 내리는 것일까 정말 너는 잃어버린 역사 처럼 훌훌이 가는 것이나

'그리워서' 단어로 시작하는 총 24 단어의 운동주체 시 생성 예제

그리워서 가거라 많은 별빛이 내린 언덕 위에 눈을 감으면 불을 끄웁니다 가슴에 꽂고 병실 안으로 의사는 젊은이의 병을 모른다 나한테는 병이 없다고

References

- https://en.wikipedia.org/wiki/Long_short-term_memory
- <https://ko.wikipedia.org/wiki/하늘과바람과별과시>

License

MIT © 송은정