

RNN PackedSequence



PackedSequence

Padding method

Packing method

Padding과 packing의 관계

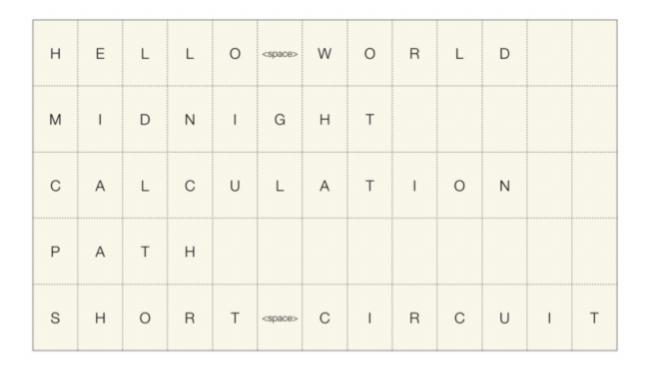
PackedSequence



길이가 각각 다른 sequence data를 하나의 batch로 묶는 방법이 있을까?

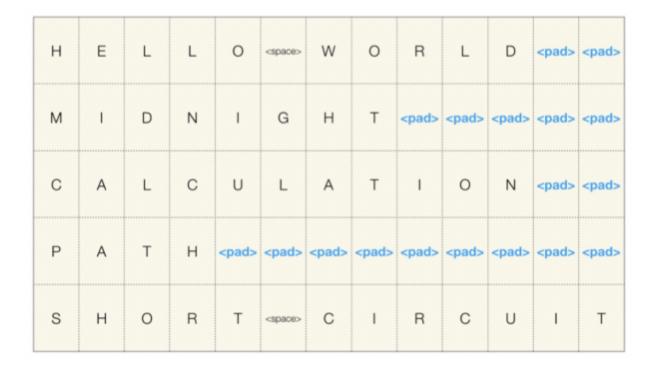
- Text
- Audio

길이가 미정인 경우가 많다. 반면 이미지 같은 경우에는 128x128처럼 크기가 정해져 있다.



이렇게 서로 다른 길이의 텍스트 시퀀스 데이터를 어떻게 다룰 수 있을까?

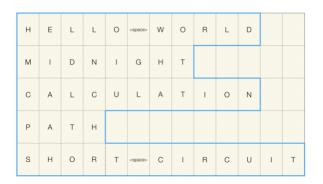
Padding method

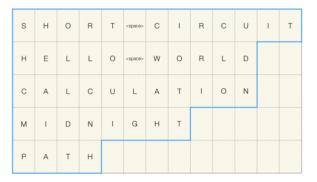


가장 긴 시퀀스의 길이에 맞추어 나머지 길이만큼을 padding로 채워준다. 이렇게 되면 batch size x 가장 긴 문자열의 길이 만큼의 크기를 가진 텐서로 다룰 수 있다.

단, 뒷부분을 padding으로 채우기 때문에 불필요한 계산이 추가적으로 수행된다는 단점이 존재한다.

Packing method





Padding을 사용하지 않고 시퀀스 길이 정보를 저장하는 방식이다. PyTorch에서 제대로 동작하기 위해서는 batch data가 길이 내림차순 순으로 정렬되어 있어야 한다. (오른쪽)

Padding 방법에 비해 효율적이고 padding을 사용하지 않는다는 장점이 있지만, 내림차순 정렬과 구현 방법이 조금 더 복잡하다는 단점이 있다.

Padding과 packing의 관계

