



RNN PackedSequence

🔍 상태	RNN
👤 담당자	

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.

H	E	L	L	O	<space>	W	O	R	L	D		
M	I	D	N	I	G	H	T					
C	A	L	C	U	L	A	T	I	O	N		
P	A	T	H									
S	H	O	R	T	<space>	C	I	R	C	U	I	T

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

Padding method

H	E	L	L	O	<space>	W	O	R	L	D	<pad>	<pad>
M	I	D	N	I	G	H	T	<pad>	<pad>	<pad>	<pad>	<pad>
C	A	L	C	U	L	A	T	I	O	N	<pad>	<pad>
P	A	T	H	<pad>	<pad>	<pad>	<pad>	<pad>	<pad>	<pad>	<pad>	<pad>
S	H	O	R	T	<space>	C	I	R	C	U	I	T

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

Diagram illustrating the placement of words in a crossword puzzle grid. The grid is 10x10. The words are placed horizontally as follows:

- Row 1: H E L L O <space> W O R L D
- Row 2: M I D N I G H T
- Row 3: C A L C U L A T I O N
- Row 4: P A T H
- Row 5: S H O R T <space> C I R C U I T

The words are: HELLO, WORLD, MIDNIGHT, CALCULATION, PATH, and SHORTCIRCUIT.

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

```
graph TD; A["List of Tensors  
[Tensor]"] -- "pack_sequence" --> B["PackedSequence"]; A -- "pad_sequence" --> C["PaddedSequence  
Tensor of shape  
[T x Batch x *] or  
[Batch x T x *]"]; B -- "pad_packed_sequence" --> C; C -- "pack_padded_sequence" --> B;
```

The diagram illustrates the relationship between three data structures:

- List of Tensors [Tensor]**: The starting point at the top.
- PackedSequence**: Reached from the List of Tensors via the `pack_sequence` operation.
- PaddedSequence**: Reached from the List of Tensors via the `pad_sequence` operation. It is a Tensor of shape $[T \times \text{Batch} \times *]$ or $[\text{Batch} \times T \times *]$.

Operations between PackedSequence and PaddedSequence:

- `pad_packed_sequence`: Converts a PackedSequence into a PaddedSequence.
- `pack_padded_sequence`: Converts a PaddedSequence back into a PackedSequence.

T is maximum sequence length

