

[순환 신경망(RNN)]

: 입력과 출력을 시퀀스로 처리하는 모델

1. 문맥을 이용하여 모델 만들기

(1) 토큰화와 정수 인코딩

```
from keras_preprocessing import Tokenizer # 토큰화를 위해 필요한 패키지 import
t = Tokenizer( ) # Tokenizer 객체 생성
t.fit_on_texts([ ]) # Tokenizing
encoded = t.texts_to_sequences([ ]) # 텍스트의 순서 생성 (정수 인코딩)
t.word_index( ) # Tokenizing한 각 단어들(key)과
# 정수인코딩된 수(value)를 딕셔너리로 출력
```

(2) pad_sequences(): 모든 데이터의 길이를 맞춰준다.

```
from keras_preprocessing.sequence import pad_sequences # 패키지 import
sequences = pad_sequences(sequences, maxlen=6, padding='pre')
# maxlen = ; 시퀀스의 길이 설정 /
# padding = 'pre' ; 자릿수가 maxlen에서 설정한 값보다 적을 경우 앞에 0을 붙여라
```

(3) 모델 생성

```
ex) from keras_preprocessing.layers import SimpleRNN, Dense, Embedding # 패키지 import
model.add(SimpleRNN(32)) # 히소행렬로 변환
```