



8장 어텐션

윤 예 준

INDEX

01 어텐션 구조

02 어텐션을 갖춘 seq2seq 구현

03 어텐션 평가

04 어텐션에 관한 남은 이야기

05 어텐션 응용

01

RNN의 문제점

기울기 소실과 기울기 폭발 원인

02

기울기 소실과 LSTM

LSTM 계층 조립

게이트

전용 가중치 매개변수 이용

가중치 매개변수는 학습 데이터로부터 갱신

게이트 열림 상태를 구할 때는 시그모이드 함수 사용

※ ‘게이트를 얼마나 열까’ 라는 것도 데이터로부터 (자동으로) 학습

03

LSTM 구현

LSTM 구현

04

LSTM을 사용한 언어 모델

LSTM을 사용한 언어 모델

1. 기울기 클리핑 수행
2. 데이터 크기 때문에 모든 에폭에서 평가하지 않고, 20번 반복 될 때마다 평가
3. 학습이 끝난 후 테스트 데이터를 사용해 퍼플렉서티를 평가 (LSTM의 은닉 상태와 기억셀을 재설정하여 평가 수행)
4. 학습이 완료된 매개변수들을 파일로 저장

05

RNNLM 추가 개선

RNNLM 추가 개선 드롭아웃

참고: A theoretically grounded application of dropout in recurrent neural networks

04

정 리

GRU

RNN 학습 문제점 : 기울기 소실, 폭발이 있음

LSTM은 output, forget, input 3개의 게이트로 이루어짐

게이트에는 전용 가중치가 있으며, 시그모이드 함수를 사용한다.

언어 모델 개선에는 LSTM 다층화, 드롭아웃, 가중치 공유 등의 기법이 효과적

THE

END

감 사 합 니 다
