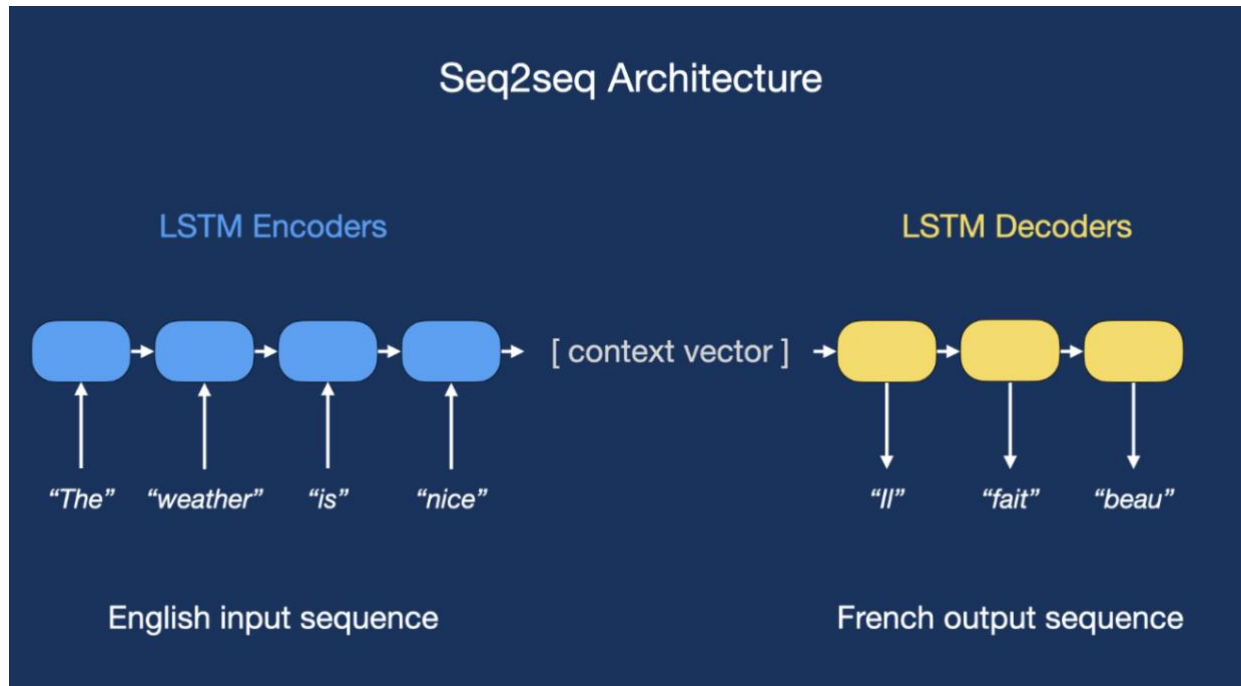


Seq2Seq

Sequence to Sequence Learning with Neural Networks (NIPS 2014)



<https://medium.com/one-minute-machine-learning/sequence-to-sequence-learning-with-neural-networks-2014-one-minute-summary-bce5e24c5e0c>

Introduction

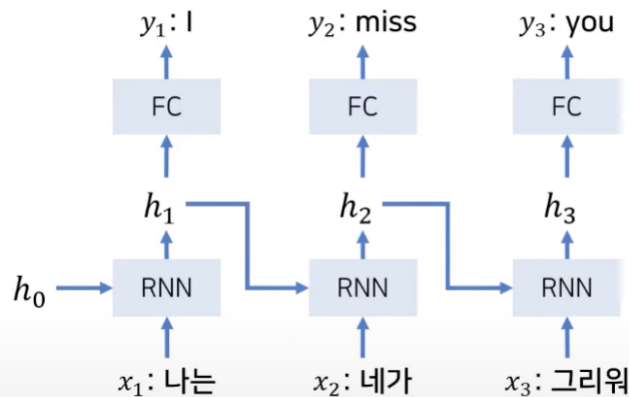
Seq2Seq 는 2014년 인공지능 분야 최우수 학회 중 하나인 NIPS에서 소개되었으며, 현대의 기계 번역 (machine translation) 기술들이 딥러닝 기반으로 전환되도록 해준 돌파구와 같은 역할을 해준 논문

- 본 논문에서는 LSTM 을 활용한 효율적인 Seq2Seq 기계 번역 아키텍처를 제안
- 아직도 많은 논문에서 베이스라인으로 비교되고 있을 정도로 중요한 논문
- Transformer(2017) 가 나오기 전까지 state-of-the-art 로 사용됨
- 2021. 12 월 기준, 약 17,000 회 인용
- Seq2Seq 는 입력된 시퀀스로부터 다른 도메인의 시퀀스를 출력하는 다양한 분야에서 활용가능하며, 챗봇과 기계 번역, 내용 요약, STT(Speech to Text) 등에서 쓰일 수 있다

등장 배경

전통적인 RNN 기반의 번역 과정

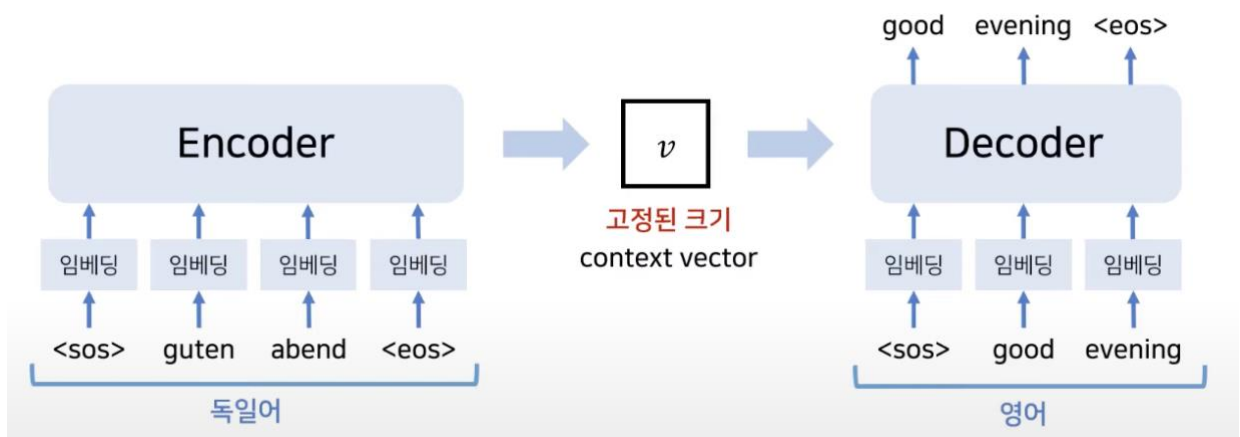
- 입력: (x_1, \dots, x_T)
- 출력: (y_1, \dots, y_T)
 - $h_t = \text{sigmoid}(W^{hx}x_t + W^{hh}h_{t-1})$
 - $y_t = W^{yh}h_t$



<https://medium.com/one-minute-machine-learning/sequence-to-sequence-learning-with-neural-networks-2014-one-minute-summary-bce5e24c5e0c>

- 전통적인 RNN 기반의 기계번역은 입력과 출력의 크기가 같다고 가정하기에 정확한 번역에 대한 한계 존재
- 입력을 받을 때마다 출력을 진행하는 방식으로, 국가마다 다른 언어의 어순을 반영하기에 어려움이 존재
- Seq2Seq 는 전통적인 RNN 기반 번역 과정의 한계를 보완하기 위해 등장
- Seq2Seq 는 Encoder 에서 전체 문장의 정보를 담고 있는 하나의 context vector 를 추출한 이후, Decoder 를 이용하여 번역을 진행하는 방식으로 위의 한계를 보완
- 본 Seq2Seq 논문에서는, 기본적인 RNN 보다 긴 dependency 를 처리하기 적합한 LSTM 모델을 이용하여 성능을 향상시킴

동작 원리



https://github.com/ndb796/Deep-Learning-Paper-Review-and-Practice/blob/master/lecture_notes/Seq2Seq.pdf

Encoder Part와 Decoder Part로 구성된 형태를 갖고 있으며,
Seq2Seq 는 하나의 Sequence 에서 다른 Sequence 로 번역한다는 의미를 지님

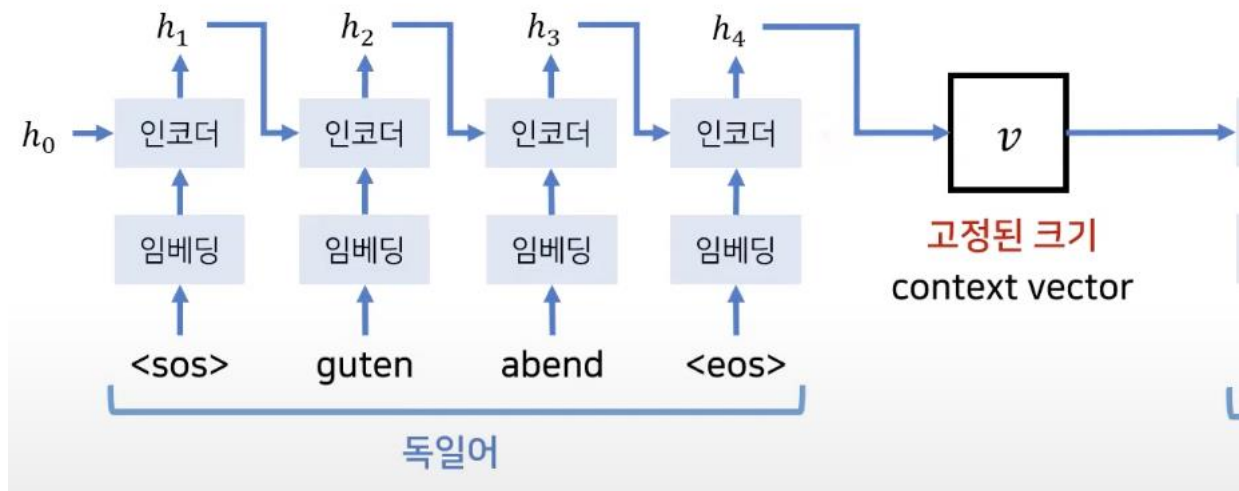
- 특정 나라의 언어로 구성된 문장이 있을 때, 다른 나라의 언어의 문장으로 번역
- Sequence 는 일반적으로 여러 개의 단어(token)로 구성된 하나의 문장을 의미
- 입력 문장이 들어왔을 때 Encoder Part 에서는 **고정된 크기**의 context vector 로 변환을 하고, 변환된 context vector 가 Decoder 로 전달된 후 적절한 문장으로 번역 되도록 architecture 를 설계
- **고정된 크기**의 context vector 에 문장의 특징을 모두 담아야 한다는 점이 특징이자 한계로 작용



https://github.com/ndb796/Deep-Learning-Paper-Review-and-Practice/blob/master/lecture_notes/Seq2Seq.pdf

- context vector 가 입력 문장에 대한 정보를 정확히 표현할 수 없다면, 번역 성능에 제한을 갖게 됨
- 이에, context vector 가 성능의 bottleneck 으로 작용

Encoder



https://github.com/ndb796/Deep-Learning-Paper-Review-and-Practice/blob/master/lecture_notes/Seq2Seq.pdf

- Encoder 는 번역하고자하는 문장이 처리되는 곳
- 번역하고자하는 문장이 순차적으로 LSTM, RNN 와 같은 Sequential 한 Data 를 처리하기에 적합한 모델에 들어가, 하나의 고정된 크기의 context vector 를 만든다
- 입력 sequence 정보 (input)를 적절히 표현할 수 있는 context vector (output)로 변환
- 모델의 hidden state 값은 단어가 들어올 때마다 갱신되며, 가장 마지막 단어까지 입력된 이후에 나온 hidden state 는 모든 문장의 정보를 포함하고 있는 문맥 vector 이기에, Encoder 의 마지막 hidden state 를 context vector 로 사용한다
- 문장의 시작을 알리기위해 sos (start-of-sentence), 문장의 끝을 알리기 위해 eos (end-of-sentence) 라는 특수 token 을 넣기도 한다

Decoder

- 종료 시점: $y_t = \langle \text{eos} \rangle$

$$p(y_1, \dots, y_{T'} | x_1, \dots, x_T) = \prod_{t=1}^{T'} p(y_t | v, y_1, \dots, y_{t-1})$$

- $v : x_1, \dots, x_T$ 에 대한 정보를 담은 벡터 표현
- y_t : 현재의 출력 단어



https://github.com/ndb796/Deep-Learning-Paper-Review-and-Practice/blob/master/lecture_notes/Seq2Seq.pdf

- context vector (input)에 담겨있는 문장의 의미를 번역하고자하는 나라의 문장 (output) 으로 변환
- context vector 를 사용하기에, 입력과 출력의 길이가 달라도 문제가 발생하지 않음
- Encoder 와 Decoder 는 유사한 Architecture 를 갖지만, 서로 다른 Parameter (가중치) 이용하여 학습하도록 만든다
- 출력의 결과로 eos 가 나오는 시기가 종료 시점이 된다
- 고정된 크기의 context vector 를 사용하기에 길이가 짧은 문장을 위주로 학습하다가 test 시에 길이가 긴 문장이 들어오게되면, 고정된 크기의 vector 가 모든 정보를 표현할 수 없어 성능 저하가 발생할 수 있다
- Seq2Seq 이후의 모델에서 위의 문제들을 많이 해결하게된다

Reference

- <https://arxiv.org/pdf/1409.3215.pdf>
- https://github.com/ndb796/Deep-Learning-Paper-Review-and-Practice/blob/master/lecture_notes/Seq2Seq.pdf
- <https://wikidocs.net/24996>
- https://yngie-c.github.io/nlp/2020/06/30/nlp_seq2seq/
- <https://medium.com/one-minute-machine-learning/sequence-to-sequence-learning-with-neural-networks-2014-one-minute-summary-bce5e24c5e0c>