

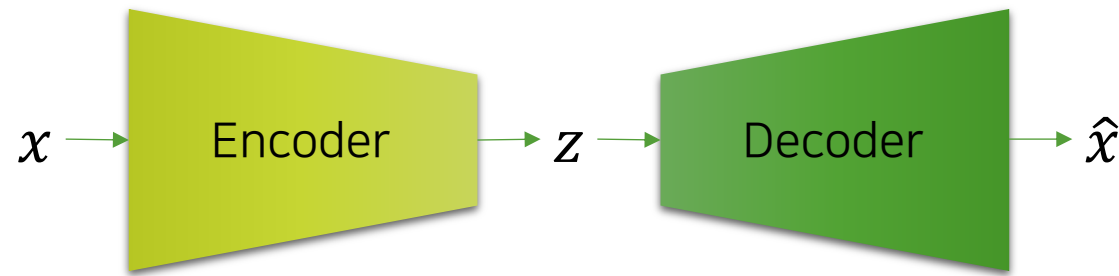
Review: NLP Introductions

Ki Hyun Kim

nlp.with.deep.learning@gmail.com

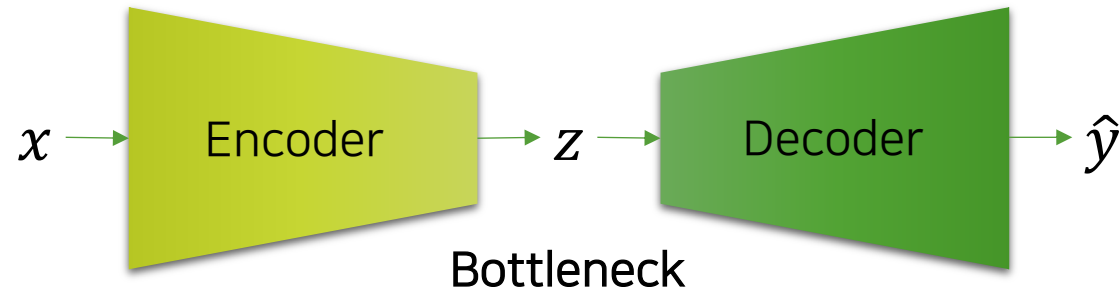
Review: Autoencoders

- 인코더(encoder)와 디코더(decoder)를 통해 압축과 해제를 실행
 - 인코더는 입력(x)의 정보를 최대한 보존하도록 손실 압축을 수행
 - 디코더는 중간 결과물(z)의 정보를 입력(x)과 같아지도록 압축 해제(복원)를 수행
- 복원을 성공적으로 하기 위해,
오토인코더(autoencoder)는 특징(feature)을 추출하는 방법을 자동으로 학습



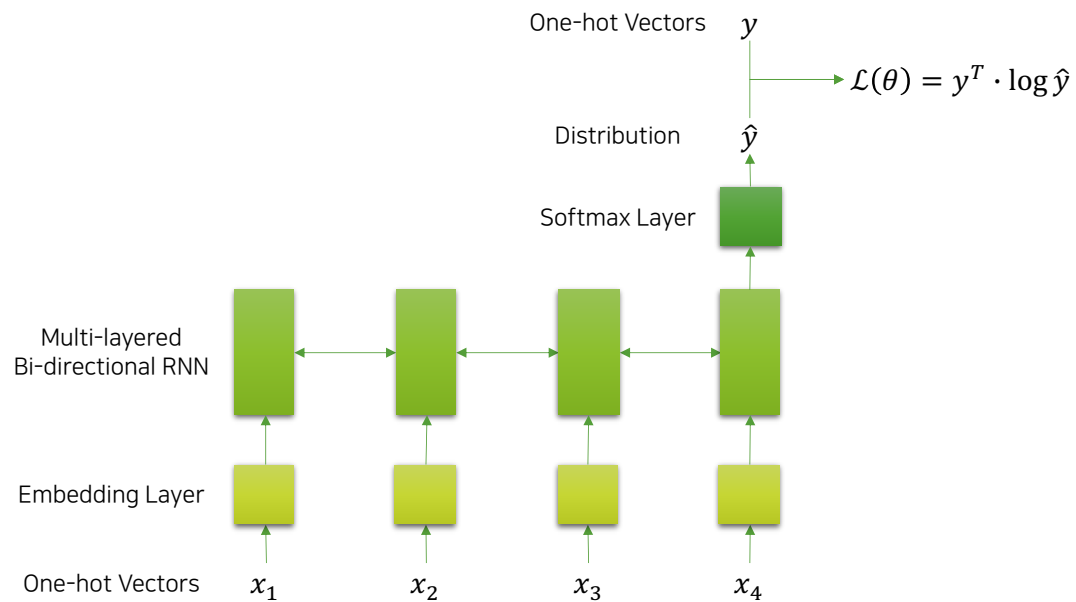
In Word2Vec

- Objective: 주어진 단어로 주변 단어를 예측하자.
- y 를 예측하기 위해 필요한 정보가 z 에 있어야 한다.
 - 주변 단어를 잘 예측하기 위해 x 를 잘 압축하자.

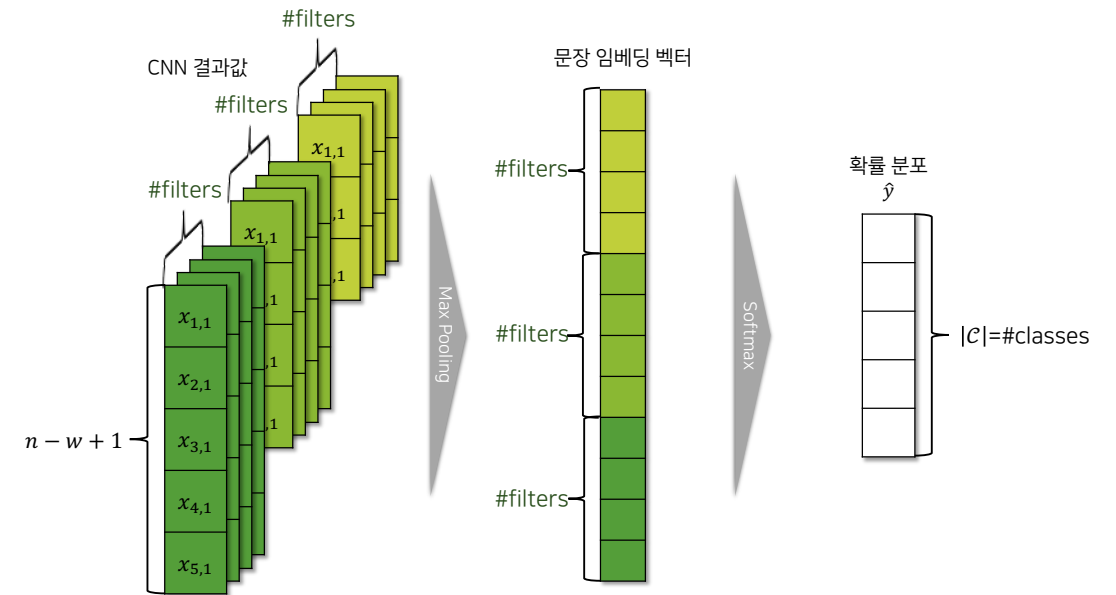


In Text Classification

Using RNN



Using CNN



Wrap-up

- 신경망은 x 와 y 사이의 관계를 학습하는 과정에서 feature를 자연스럽게 학습
 - 특히 저차원으로 축소(압축)되는 과정에서 정보의 취사/선택이 이루어짐
- Word Embedding(Skip-gram):
 - 주변 단어(y)를 예측하기 위해 필요한 정보를 현재 단어(x)에서 추출하여 압축
- Sentence Embedding(text classification):
 - Label(y)을 예측하기 위해 필요한 정보를 단어들의 시퀀스(x)로부터 추출하여 압축