

+·*~ SMARCLE PAPER REVIEW ~·*



Recent Trends in Deep Learning Based Natural Language Processing (2017)

지능기전공학부 무인이동체공학전공
송혜원

Introduction

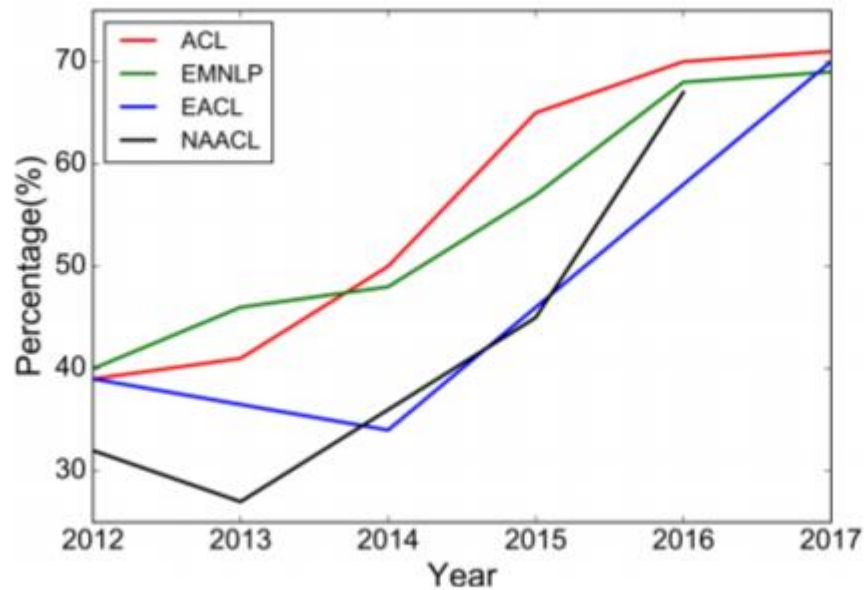


Fig. 1: Percentage of deep learning papers in ACL, EMNLP, EACL, NAACL over the last 6 years (long papers).

NLP 분야에 머신 러닝 기법 도입

Dense vector representation에 기반한 NN을 사용

다양한 NLP task에 우수한 성능



Introduction

· 머신러닝 기법이 도입되기 전에는?

Statistical 통계적 기법

데이터 100개 있다고 가정

$$P(\text{는}|\text{나}) = 0.93$$

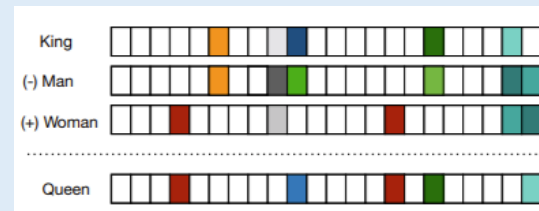
$$P(\text{밥}|\text{나는}) = 0.61$$

$$P(\text{을}|\text{나는 밥}) = 0.99$$



Neural 신경망 기반 언어 모델

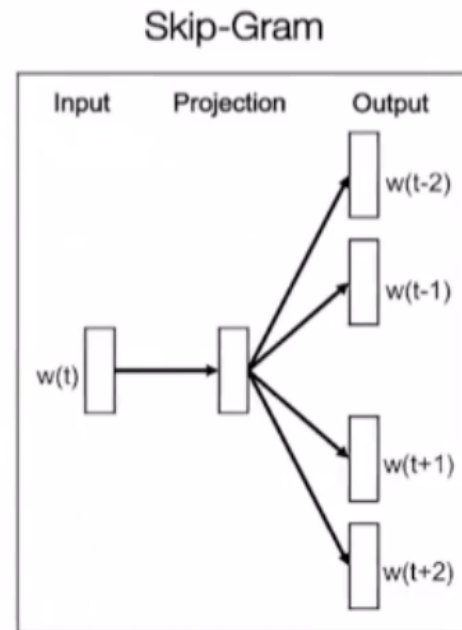
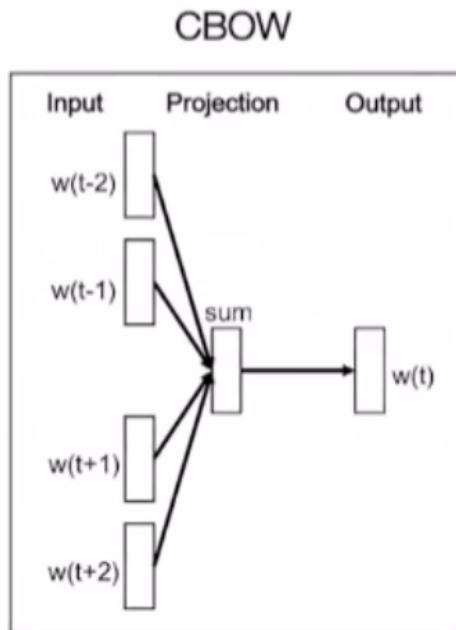
분산 표현



Model
(RNN)

Word Embedding

- Word2Vec (Hypothesis : 비슷한 분포를 가진 단어는 비슷한 의미를 가진다.)



- CBOW : 주변 -> 중심
- Skip-Gram : 중심 -> 주변

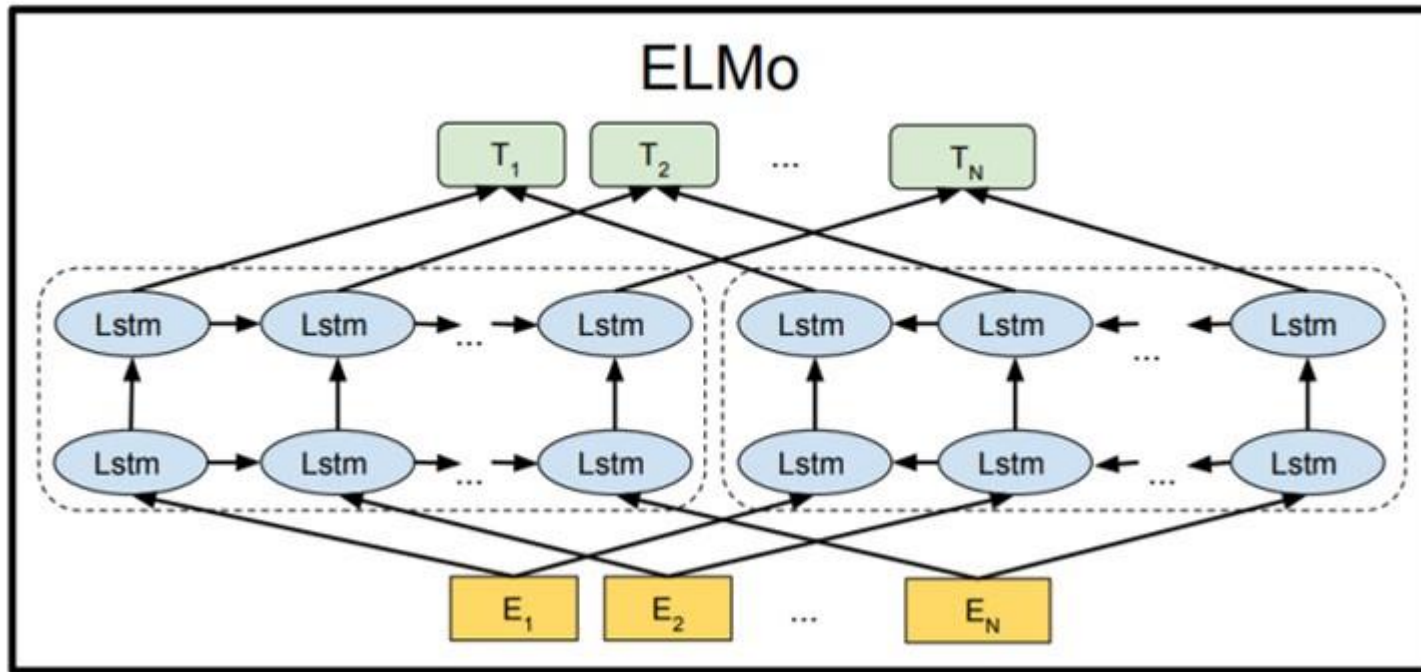
Yesterday, I ____ the exam

____, __ took ____

* Continuous Bag of Words

Word Embedding

- ELMo(Embeddings from Language Model)

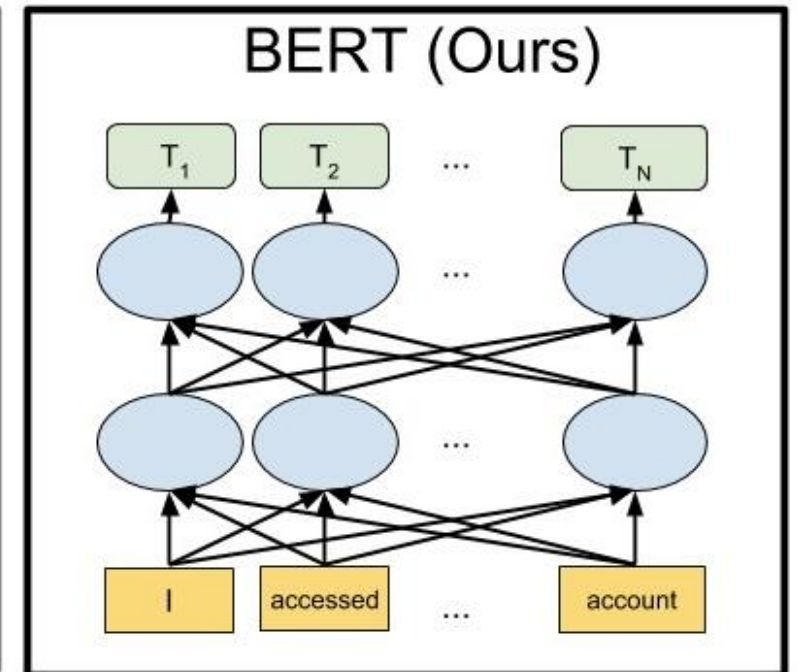
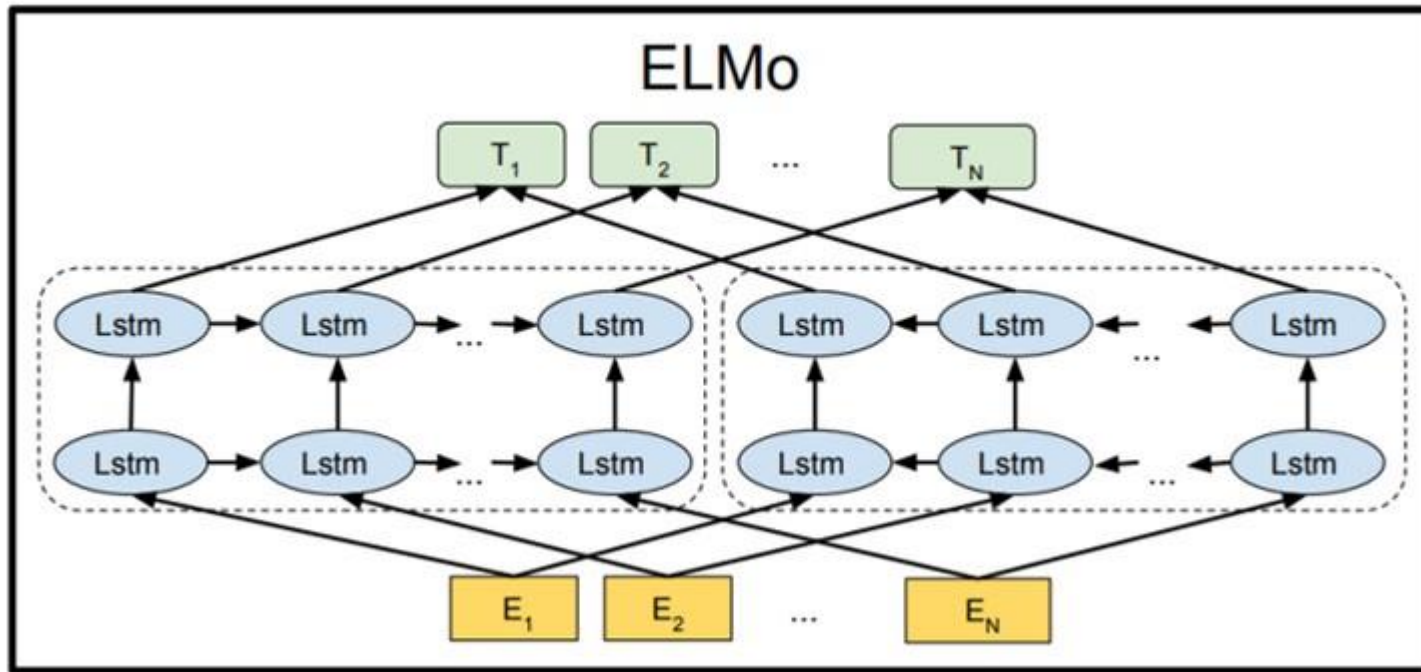


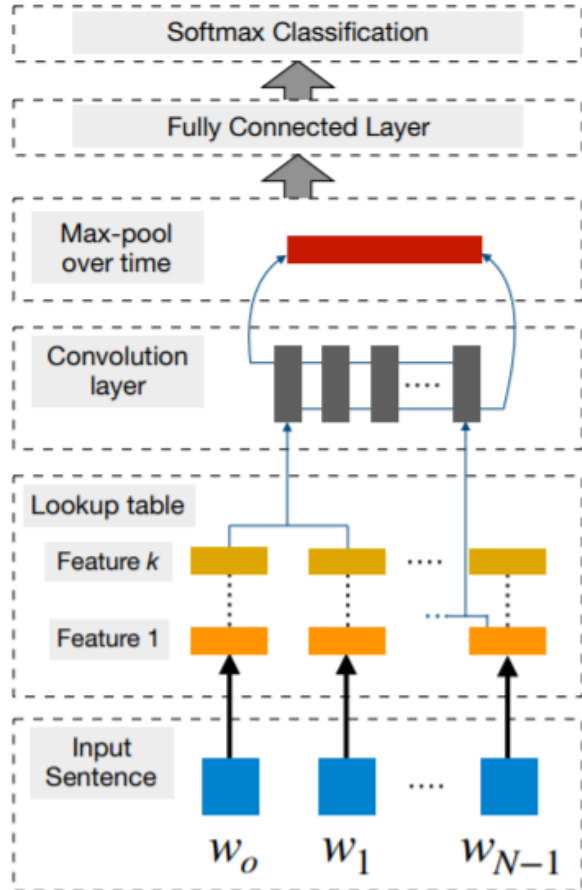
- 양방향으로 동작
- 문맥을 파악

Ex) Dog와 Hot Dog 판별

Word Embedding

- ELMo(Embeddings from Language Model)





- 단어의 등장 순서 / 문맥 정보 보존
- 합성곱, ReLU, Max pooling 과정 반복
- 학습된 결과와 비교하여 유사도 판별
- 단어 단위 예측 필요 -> window 접근법



RNN

• Simple RNN

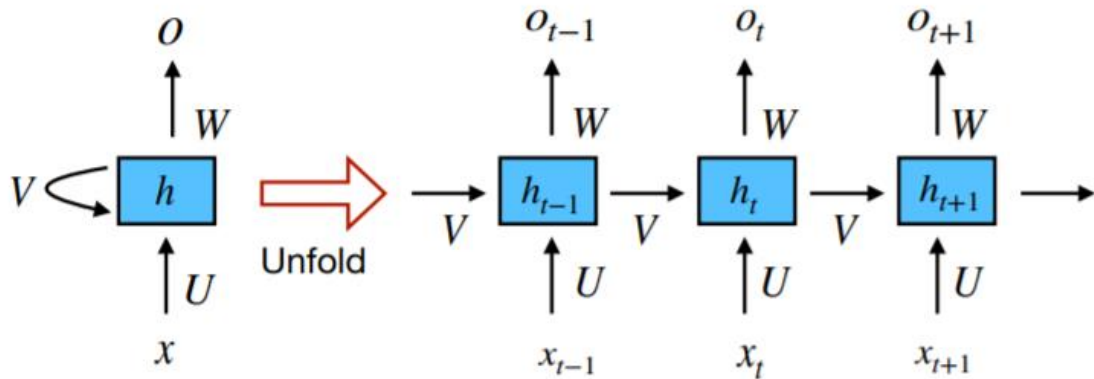
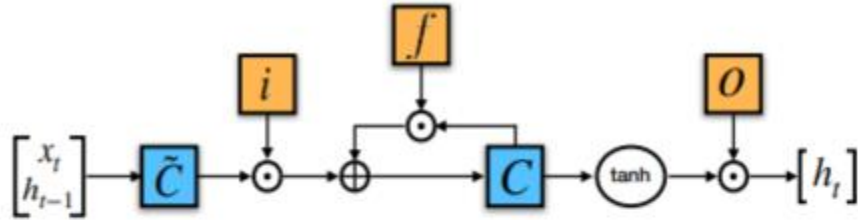


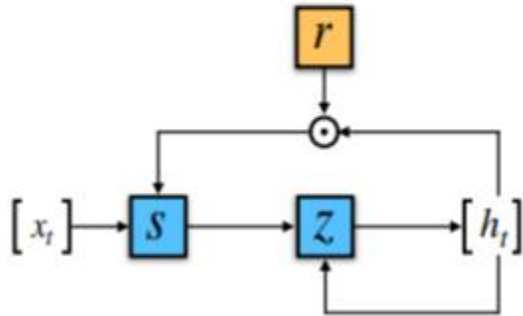
Fig. 9: Simple RNN network (Figure Source: LeCun et al. [90])

- 앞의 state가 뒤 state까지 영향
- Input = 앞 input + 앞 input*가중치
- Many to One : 감정인식
- Many to Many : 번역

- LSTM(Long Short-Term Memory) & GRU(Gated Recurrent Unit)



(1) Long Short-Term Memory



(2) Gated Recurrent Unit

- LSTM : input gate, forget gate, output gate
- GRU : update gate (input+forget) reset gate

Fig. 10: Illustration of an LSTM and GRU gate (Figure Source: Chung et al. [81])



MLM(Masked Language Model)

BERT

GPT



MLM(Masked Language Model)

GAN?

Weight?

Detail?

TBC...!