Convolutional Neural Networks for Sentence Classification

Ha-Iarm

Abstract

- Word Vector와 CNN을 활용한 문장 분류

- 이미 트레이닝 된 word vector를 활용
- Simple한 CNN 사용
- 높은 정확도

문장 분류(Sentence Classification)

- **감정 분류(Sentiment Analysis)**
- Ex)이거 정말 좋다-긍정 , 이 레스토랑 음식 정말 별로다-부정

- **주제 분류**
- EX)손흥민 MOM선정 –스포츠 , 유승민 유세 성공적-정치

Word2Vec

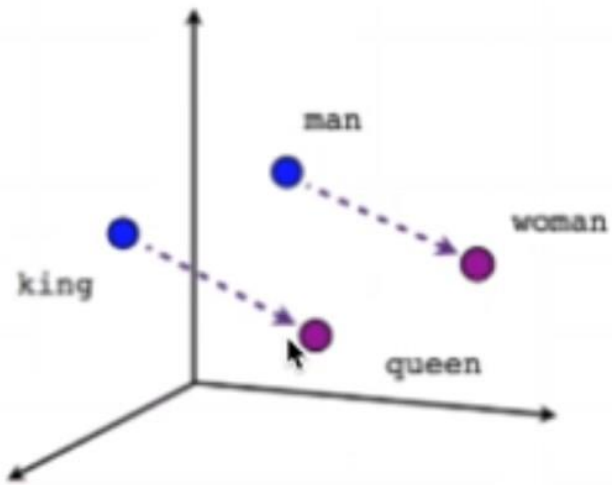
- 각 단어마다 Vector 값을 부여하자
- 단어들의 특징을 표현할 수 있도록 수치로 된 값 부여
- Ex)강아지=[2,4,5,1,3]

Word2Vec

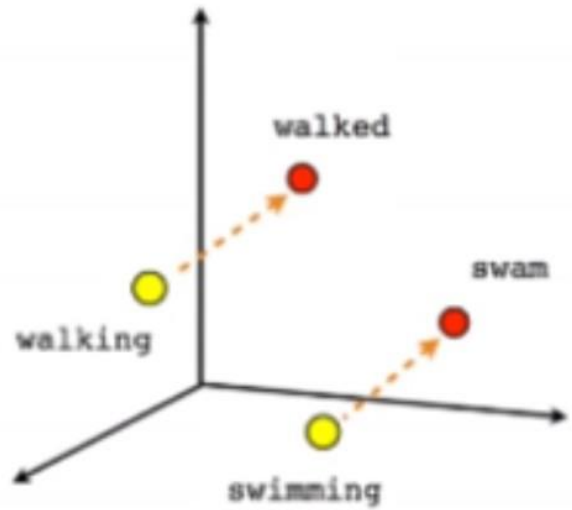
- The quick brown fox jumped over the lazy dog
  ([the,brown],quick),([quic,fox],brown),([brown,jumped],fox)

- 문장에서 나오는 **단어**들의 위치로 학습시키자
- 다양한 Window size를 이용

그랬더니 특정 방향들이 의미를 담고 있었어!



Male-Female

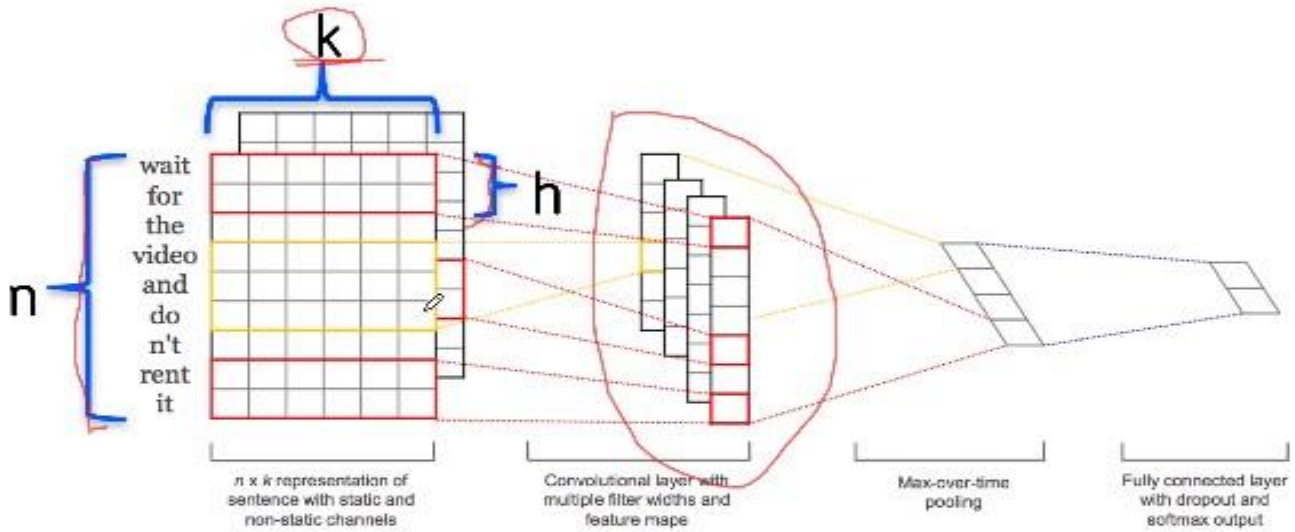


Verb tense

Country-Capital

CNN과 Word Vector를 이용한 문장 분류

Model Architecture



$$\mathbf{x}_{1:n} = \mathbf{x}_1 \oplus \mathbf{x}_2 \oplus \ldots \oplus \mathbf{x}_n, \quad (1)$$

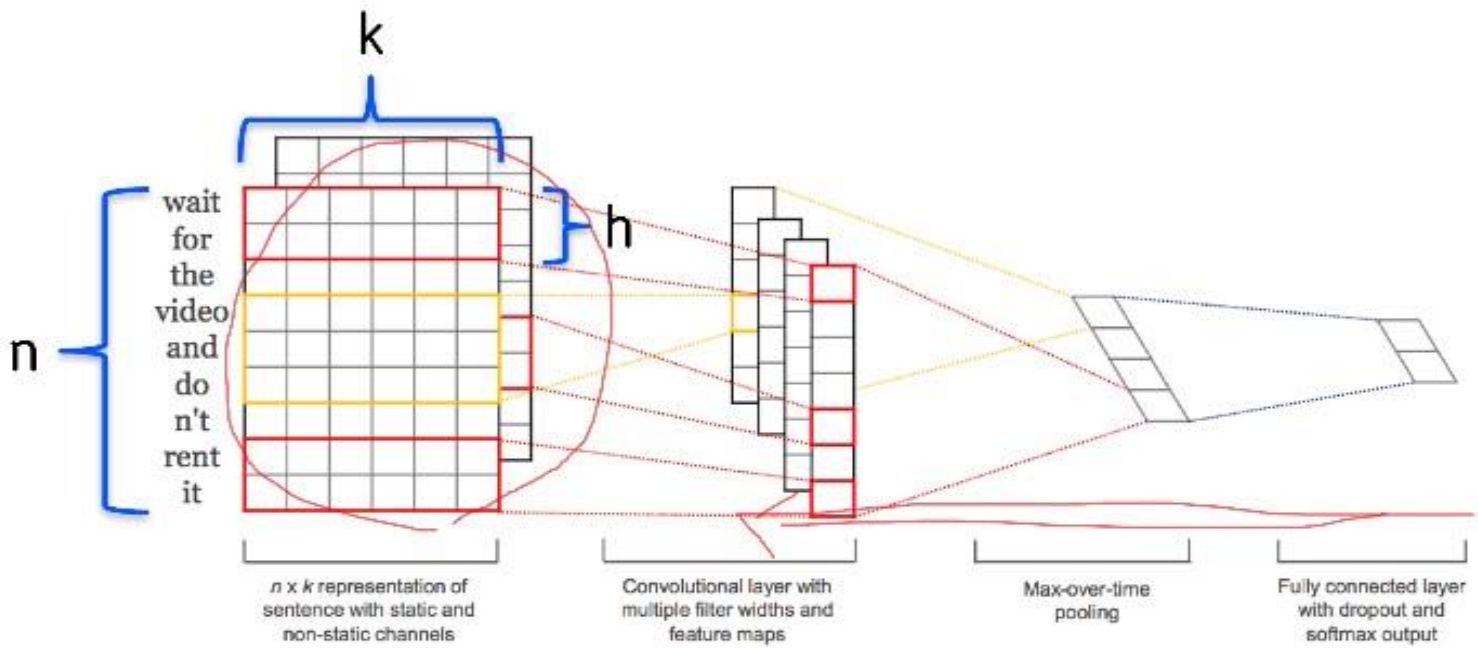
$$c_i = f(\mathbf{w} \cdot \mathbf{x}_{i:i+h-1} + b). \quad (2)$$

$$\mathbf{c} = [c_1, c_2, \ldots, c_{n-h+1}], \quad (3)$$

n : 문장에 나오는 단어의 갯수 k : Word Vector의 차원 h : 필터 윈도우 사이즈

Page2 Figure1 (Yoon Kim, Convolutional Neural Networks for Sentence Classification, EMNLP 2014)

Static, Non-static, Multichannel



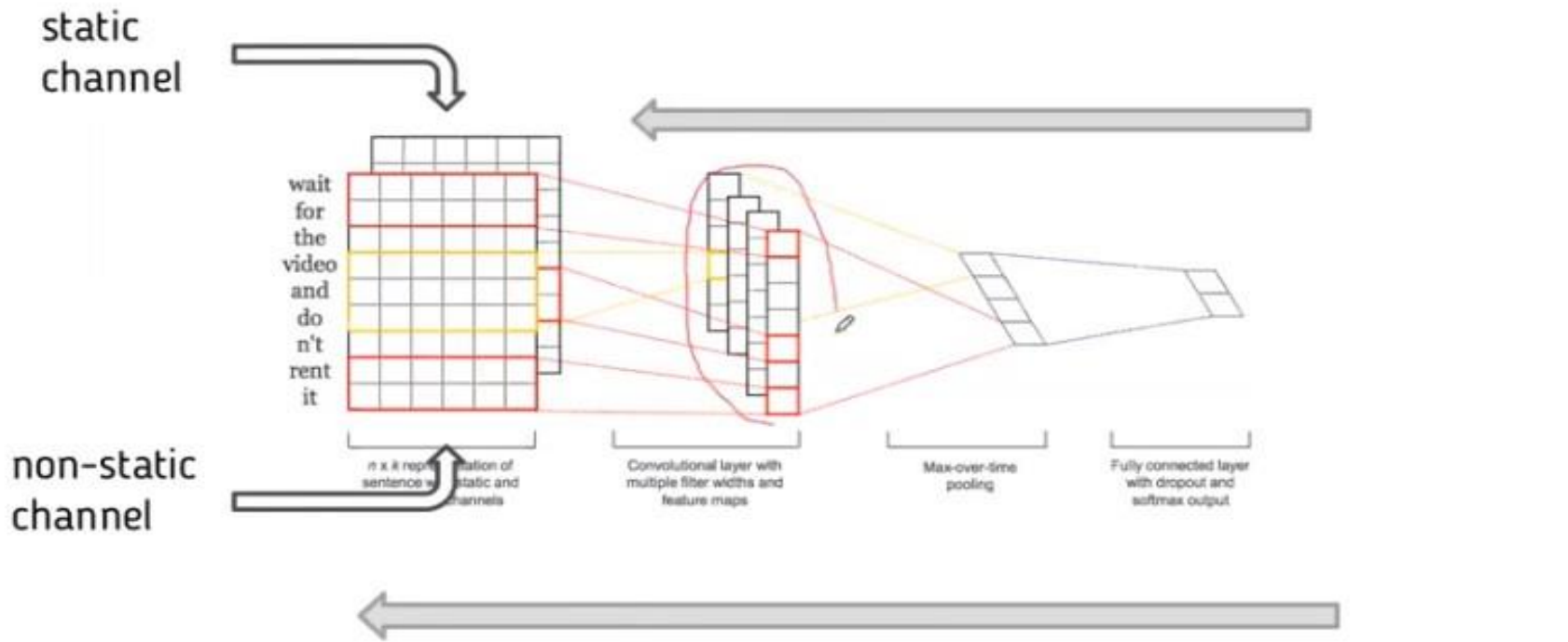
Back Propagation

static

non-static

multichannel

Multichannel



Static vs. Non-static

	Most Similar Words for	
	Static Channel	Non-static Channel
bad	*good* *terrible* *horrible* *lousy*	*terrible* *horrible* *lousy* *stupid*
good	*great* *bad* *terrific* *decent*	*nice* *decent* *solid* *terrific*
n't	*os* *ca* *ireland* *wo*	*not* *never* *nothing* *neither*
!	*2,500* *entire* *jez* *changer*	*2,500* *lush* *beautiful* *terrific*
,	*decasia* *abysmally* *demise* *valiant*	*but* *dragon* *a* *and*

Non-static으로 학습시키니 word vector가 의미를 더 잘 이해하게 되었군!

Page5 Table3 (Yoon Kim, Convolutional Neural Networks for Sentence Classification, EMNLP 2014)

Model	MR	SST-1	SST-2	Subj	TREC	CR	MPQA
CNN-rand	76.1	45.0	82.7	89.6	91.2	79.8	83.4
CNN-static	81.0	45.5	86.8	93.0	92.8	84.7	**89.6**
CNN-non-static	**81.5**	48.0	87.2	93.4	93.6	84.3	89.5
CNN-multichannel	81.1	47.4	**88.1**	93.2	92.2	**85.0**	89.4
RAE (Socher et al., 2011)	77.7	43.2	82.4	–	–	–	86.4
MV-RNN (Socher et al., 2012)	79.0	44.4	82.9	–	–	–	–
RNTN (Socher et al., 2013)	–	45.7	85.4	–	–	–	–
DCNN (Kalchbrenner et al., 2014)	–	48.5	86.8	–	93.0	–	–
Paragraph-Vec (Le and Mikolov, 2014)	–	**48.7**	87.8	–	–	–	–
CCAE (Hermann and Blunsom, 2013)	77.8	–	–	–	–	–	87.2
Sent-Parser (Dong et al., 2014)	79.5	–	–	–	–	–	86.3
NBSVM (Wang and Manning, 2012)	79.4	–	–	93.2	–	81.8	86.3
MNB (Wang and Manning, 2012)	79.0	–	–	**93.6**	–	80.0	86.3
G-Dropout (Wang and Manning, 2013)	79.0	–	–	93.4	–	82.1	86.1
F-Dropout (Wang and Manning, 2013)	79.1	–	–	**93.6**	–	81.9	86.3
Tree-CRF (Nakagawa et al., 2010)	77.3	–	–	–	–	81.4	86.1
CRF-PR (Yang and Cardie, 2014)	–	–	–	–	–	82.7	–
SVM_S (Silva et al., 2011)	–	–	–	–	**95.0**	–	–

Q & A

Thank you!

Paper:Convolutional Neural Networks for Sentence Classification