

2021학년도 2학기 언어와 컴퓨터

제24강 벡터 의미론 (2)

박수지

서울대학교 인문대학 언어학과

2021년 12월 8일 수요일

오늘의 목표

- 1 Word2Vec 모형의 활용 방법과 한계를 설명할 수 있다.

Word2Vec

Skip-gram with negative sampling

대상 단어 t , 문맥 단어 c , 노이즈 단어 n_i ($i = 1, 2$)가 주어졌을 때

$$\begin{aligned}\text{손실함수 } L_{CE} &= - \left[\log P(+|t, c) + \sum_{i=1}^2 \log P(-|t, n_i) \right] \\ &= - \left[\log \frac{1}{1 + e^{-t \cdot c}} + \sum_{i=1}^2 \log \frac{1}{1 + e^{t \cdot n_i}} \right]\end{aligned}$$

비교

로지스틱 회귀분석 분류기 가중치 a 는 변수, 데이터 x 는 상수

Word2Vec (Skip-gram) 대상 벡터 t 와 문맥 c 모두 변수

Word2Vec

Skip-gram with negative sampling

대상 단어 어휘 V 의 i 번째 단어

문맥 단어 어휘 V 의 j 번째 단어

$$T \times C = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1d} \\ \dots & \dots & \dots & \dots \\ t_{i1} & t_{i2} & \dots & t_{id} \\ \dots & \dots & \dots & \dots \\ t_{|V|1} & t_{|V|2} & \dots & t_{|V|d} \end{bmatrix} \cdot \begin{bmatrix} c_{11} & \vdots & c_{1j} & \vdots & c_{1|V|} \\ c_{21} & \vdots & c_{2j} & \vdots & c_{2|V|} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_{d1} & \vdots & c_{dj} & \vdots & c_{d|V|} \end{bmatrix}$$

임베딩 선택 방법

- 1 $t_i = [t_{i1}, \dots, t_{id}]$
- 2 $t_i + c_i = [t_{i1} + c_{i1}, \dots, t_{id} + c_{id}]$
- 3 $t_i \oplus c_i = [t_{i1}, \dots, t_{id}, c_{i1}, \dots, c_{id}]$

모형 매개변수

- d (벡터의 차원 수)
- L (window size — 이웃의 기준)

유사한/연관된 단어 찾기

단어의 의미 관계

연관성 두 단어가 서로 이웃인 경우 (syntagmatic association)

- 예: 연필-공책

유사성 두 단어의 이웃이 비슷한 경우 (paradigmatic association)

- 예: 연필-볼펜

유사한/연관된 단어 찾기

Window size와 단어의 의미 관계

“Hogwarts”와 가장 “가까운” 단어는? (Levy and Goldberg 2014)

Window size ± 2 유사한 단어 (2차 공기, paradigmatic association)

- Sunnydale (from Buffy the Vampire Slayer)
- Evernight (from a vampire series)

Window size ± 5 연관된 단어 (1차 공기, syntagmatic association)

- Dumbledore
- Malfoy
- half-blood

Word2Vec과 의미 변화

시대에 따라 단어의 ‘이웃’이 바뀌는 것을 볼 수 있다.

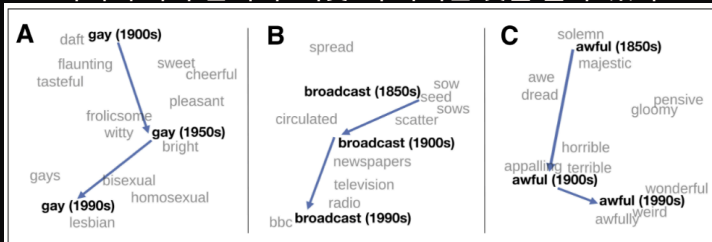


Figure 6.17 A t-SNE visualization of the semantic change of 3 words in English using word2vec vectors. The modern sense of each word, and the grey context words, are computed from the most recent (modern) time-point embedding space. Earlier points are computed from earlier historical embedding spaces. The visualizations show the changes in the word *gay* from meanings related to “cheerful” or “frolicsome” to referring to homosexuality, the development of the modern “transmission” sense of *broadcast* from its original sense of sowing seeds, and the pejoration of the word *awful* as it shifted from meaning “full of awe” to meaning “terrible or appalling” (Hamilton et al., 2016).

Word2Vec과 편향

father : doctor = mother : X \Rightarrow X = nurse??

편향의 문제

- 인종·성 등에 관한 편견을 단순히 반영할 뿐만 아니라 확대하고 재생산한다.
- 실제로 유색 인종의 이력서가 부정적으로 평가되는 등 사회적 문제를 일으킨다.

남은 할 일

코퍼스에서 Word2Vec 모델을 훈련시키기

실습 코드:

<https://colab.research.google.com/drive/1FxaqyGrBVHfFl77XQPrYyFuxCkT6Ao>

[언어와 컴퓨터]에서 배운 것

SLP Chs. 2-6

- 2 정규표현식: 코퍼스에서 단어의 패턴을 검색할 수 있다.
- 3 N그램 언어 모형: 코퍼스로부터 문장의 확률을 구하고 문장을 생성할 수 있다.
 - 문장 자동 완성
- 4 단순 베이지 분류기: 문서의 단어 빈도를 사용하여 문서의 범주를 예측할 수 있다.
 - 감정 분석
- 5 로지스틱 회귀분석: 문서의 다양한 특성값을 사용하여 문서의 범주를 예측할 수 있다.
 - TF-IDF
- 6 벡터 의미론: 단어의 의미를 벡터공간의 고정된 점으로 표현할 수 있다.
 - Word2Vec

[컴퓨터언어학]에서 배울 것

SLP Chs. 7-11

- 7 신경망 언어 모형: 코퍼스로부터 문장의 확률을 구하고 문장을 생성할 수 있다.
- 8 은닉 마르코프 모형: 단어의 연쇄로 이루어진 데이터에 태그를 부여할 수 있다.
 - 품사 태깅
- 9 순환 신경망: 단어의 연쇄로부터 문서의 범주를 예측할 수 있다.
- 10 seq2seq: 출발어 단어의 연쇄를 받아서 도착어 단어의 연쇄를 내보낼 수 있다.
 - 기계 번역
 - 문서 요약
- 11 문맥화된 단어 임베딩: 문맥에 따라 달라지는 단어의 의미를 표현할 수 있다.
 - BERT
 - GPT-2(3)