

언어모델이란?

- 자연어의 통계적 패턴
 - 문장에 어떤 단어가 (많이) 쓰였는가 : 백오브워즈 가정
 - 단어가 어떤 순서로 등장하는가 : 언어모델
 - 어떤 단어가 같이 나타났는가 : 분포 가정

언어모델이란?

- 딥러닝의 발전 이전에도 있었던 개념
- 언어를 모델링하고자 **단어 시퀀스에 확률을 부여**하는 모델이다.
- 잘 학습된 언어모델은 어떤 문장이 더 “자연스러운지”, 또한 주어진 시퀀스 다음에는 무엇이 오는게 자연스러운지를 알 수 있다.
- 단어가 N개 주어진 상황에서 언어모델은 N개 단어가 동시에 나타날 확률, 즉 $P(w_1, w_2, w_3 \dots w_n)$ 을 반환합니다.