

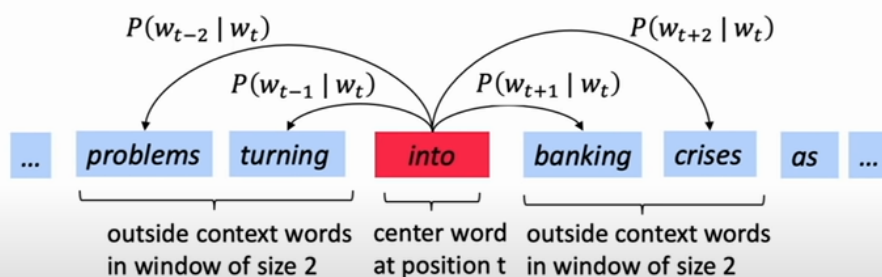
Week 1

☰ 태그	
☰ 열	

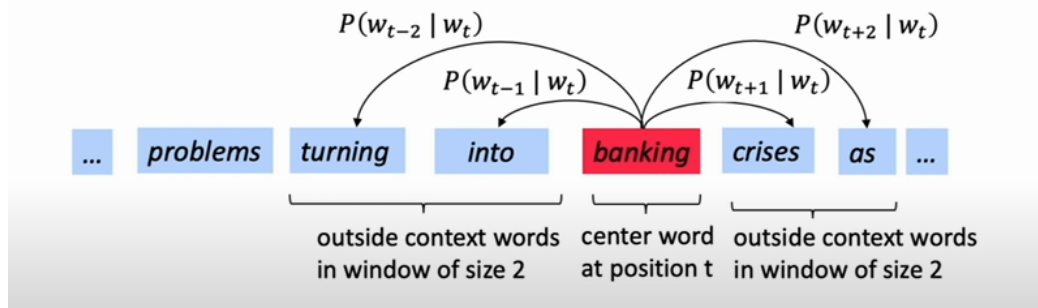
Lecture 1 – Introduction and Word Vectors

- 단어는 one-hot vector로 나타낼 수 있다.
 - Ex) motel = [0 0 0 0 1 0 0 0]
- 벡터 그 자체만으로는 단어 간에 어떤 연관성이 있는지 알 수 없다.
- 그럼 단어 간의 관계는 어떻게 나타내는가?
 - Solutions
 - Distributional semantic
 - When a word w appears in a text, its **context** is the set of words that appear nearby (within a fixed-size window).
 - Numeric vector로 표현하는 것은 동일. 그러나 모든 수가 0이 아닌 수로 표현되는 dense vector로 표현.
 - Word2Vec

- Example windows and process for computing $P(w_{t+j} | w_t)$



- Example windows and process for computing $P(w_{t+j} | w_t)$



- objective function의 값을 최소화할수록 예측 정확도가 높아진다.
- 여기서 예측이란 center word가 정해졌을 때 정해진 window 사이즈 내의 context words를 예측하는 것을 말한다.
- 단어별로 두 가지의 벡터를 이용해서 context words의 likelihood를 계산한다.
- softmax function은 어떤 숫자든 확률 분포로 매칭해준다.
- Training model
 - 모든 vector gradient를 계산해서 모델을 학습시킨다.
 - 어떤 단어가 가장 높은 확률로 올 수 있는지 계산하는 것이 목표.