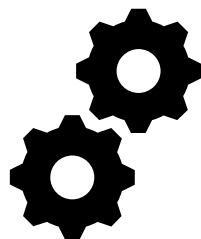


Chapter 04. 자연어처리 (Natural Language Processing)

정량 지표

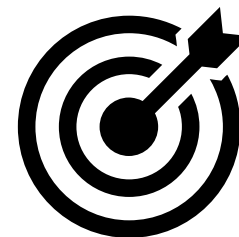
손실 함수 vs. 성능 척도

손실 함수 (Loss function)



- 최적화 이론에서 최소화 하고자 하는 함수
- 알고리즘 학습 중 아직 ‘얼마나 못 하는지’ 표현
- 보통 미분 가능한 함수를 사용
- 목적 함수, 비용 함수, 에너지 함수 등으로 불림
- 학습 과정에서 중요한 역할을 함

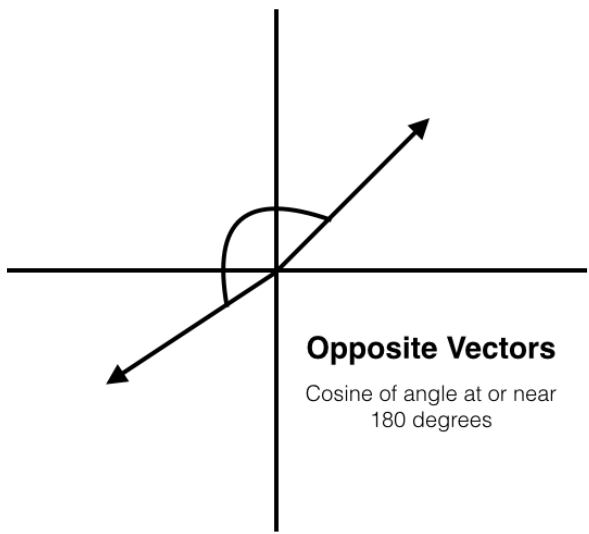
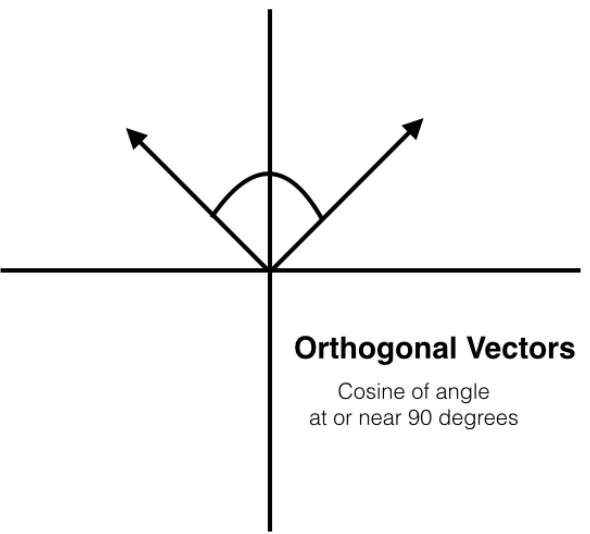
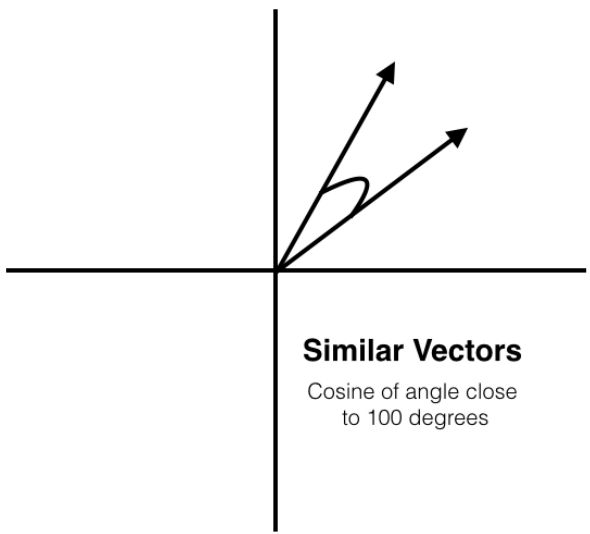
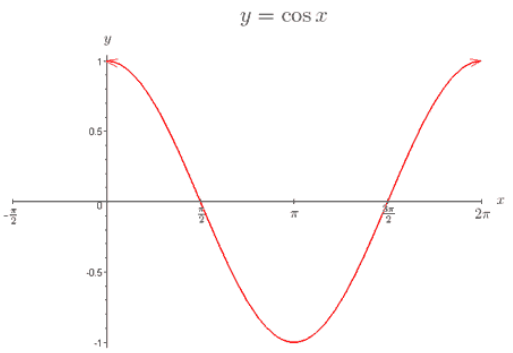
성능 척도 (Performance measure)



- 학습된 알고리즘의 성능을 측정하는 지표
- 정량적으로 알고리즘을 비교/평가하기 위함
- 미분 가능 여부에 관계 없이 사용
- 학습 과정에서는 사용되지 않음
- ‘비즈니스 목표’와 연관이 높음

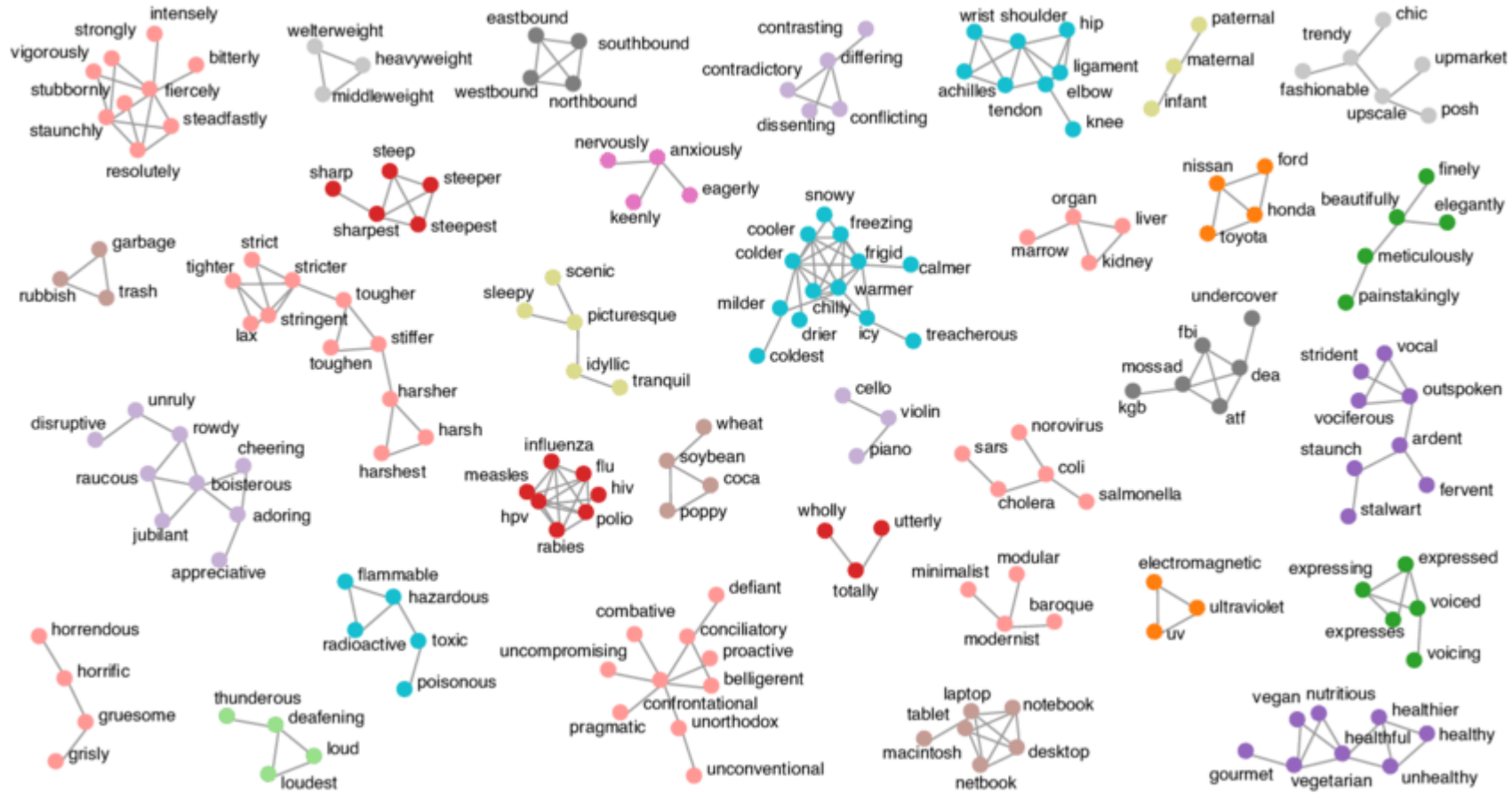
Cosine Similarity

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}},$$



두 ‘의미’를 가진 벡터를 비교할 때 유용한 Cosine Similarity

Cosine Similarity – Application



기계 번역기의 성능 평가

문제: 나는 고양이가 정말 좋고 세 마리를 키우고 있어요.

번역1:

I like cats so I have three

번역2:

I really like cats and I have three.

정답: I really like cats and live with three of them.

구글 번역기(좌), 네이버 파파고(우)의 성능을 비교해 보자.

Unigram Precision

$$\text{Unigram Precision} = \frac{\text{Ca에 나온 단어 중 Ref에 있는 단어의 수}}{\text{Ca에 있는 단어의 수}}$$

정답: I really like cats and live with three of them.

번역1:

I like cats so I have three

$$\frac{5}{7}$$

번역2:

I really like cats and I have three.

$$\frac{7}{8}$$

Unigram Precision – Bad case

$$\text{Unigram Precision} = \frac{\text{Ca에 나온 단어 중 Ref에 있는 단어의 수}}{\text{Ca에 있는 단어의 수}}$$

정답: I really like cats and live with three of them.

번역?:
Cats cats cats cats cats cats cats cats cats cats

$$\frac{10}{10}$$

Modified Unigram Precision

$$\text{Modified Unigram Precision} = \frac{\text{Ca에 있는 단어 중 Ref에 있는 단어의 수} - \text{Ref를 넘어서는 횟수}}{\text{Ca에 있는 단어의 수}}$$

정답: I really like cats and live with three of them.

번역1:

I like cats so I have three

$\frac{5}{7}$

번역2:

I really like cats and I have three.

$\frac{7}{8}$

Modified Unigram Precision

$$\text{Modified Unigram Precision} = \frac{Ca \text{에 있는 단어 중 } Ref \text{에 있는 단어의 수} - Ref \text{를 넘어서는 횟수}}{Ca \text{에 있는 단어의 수}}$$

정답: I really like cats and live with three of them.

번역?:

Cats cats cats cats cats cats cats cats cats cats

$$\frac{1}{10}$$

N-Gram

정답: I really like cats and live with three of them.

Bigram	I really, really like, like cats, cats and, and live, live with, with three, three of, of them
Trigram	I really like, really like cats, like cats and, cats and live, and live with, live with three, with three of, three of them
4-gram	I really like cats, really like cats and, like cats and live, cats and live with, and live with three, live with three of, with three of them

N-Gram Precision

$$p_n = \frac{\text{Ca에 나온 단어 중 Ref에 있는 } n - \text{Gram의 수(Clipped)}}{\text{Ca에 있는 } n - \text{Gram의 수}}$$

정답: I really like cats and live with three of them.

번역1:

I like cats so I have three

I like, like cats, cats so, so I, I have, have three

$\frac{1}{6}$

번역2:

I really like cats and I have three.

I really, really like, like cats, cats and, and I,
I have, have three

$\frac{5}{7}$

BLEU Bilingual Evaluation Understudy

$$BLEU = BP \times \exp \left(\sum_{n=1}^N w_n \log p_n \right)$$

$$BP = \begin{cases} 1, & \text{if } c > r \\ \exp(1 - r/c), & \text{if } c \leq r \end{cases} \quad \text{Brevity Penalty: 길이가 짧을 경우 페널티}$$

N개의 n -Gram Precision을 합산하는 형태이다. 단순히 길이가 짧아서 점수가 높아지는 것을 방지한다.

BLEU Bilingual Evaluation Understudy

정답: I really like cats and live with three of them.

번역1:
I like cats so I have three

번역2:
I really like cats and I have three.

지표	번역1	번역2
p1	5/7	7/8
p2	1/6	5/7
p3	0	3/6
p4	0	2/5
BP	0.65	0.78
BLEU	0.143	0.485