

AI 개발 실무

11. 텍스트 분류

김 윤 기 교수



11 week

A I 개 발 실 무 | 김 윤 기

텍스트 분류

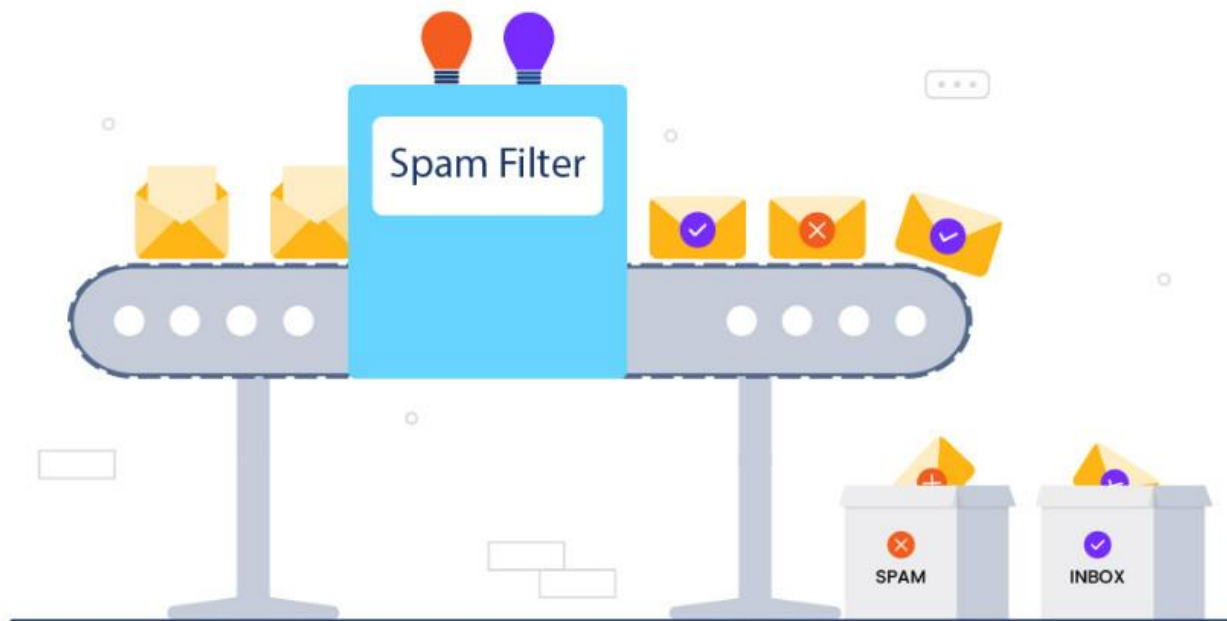


- » 베이지 정리를 이해할 수 있다.
- » 나이브 베이지 분류를 이해할 수 있다.
- » 베이저안 필터를 구현할 수 있다.

① 텍스트 분류

② 텍스트 분류 실습

스팸 메일을 어떻게 구분할까?



키워드 기반 스팸 분류가 아닌,
데이터를 학습시켜 자동으로 스팸을 분류

CHAPTER

01

텍스트 분류

1. 텍스트 분류란?

주어진 텍스트를 미리 정의된 카테고리 중 하나로 자동 분류하는 과정

주로 지도 학습(supervised learning) 방식을 사용하여 이전에 분류된 텍스트 데이터를 사용하여 새로운 데이터를 자동으로 분류

스팸 메일 분류, 뉴스 카테고리 분류, 감성 분석 등의 문제를 처리

텍스트에서 단어의 빈도(word frequency),
TF-IDF(Term Frequency-Inverse Document Frequency)
등으로 특징을 추출하여, 해당 특징으로 분류 모델을 학습

데이터 수집을 통해 확보된 비정형 데이터를 분류할 수 있음

2. 텍스트 분류 모델의 종류

① 나이브 베이즈(Naïve Bayes)

- » 베이즈 정리(Bayes' theorem)를 이용한 확률 기반 분류 알고리즘
- » 각 클래스(class)에 속하는 특징(feature)들의 조건부 확률(conditional probability)을 계산하여 가장 높은 확률을 갖는 클래스로 분류하는 방식
- » 각 특징들은 서로 독립적이라고 가정

2. 텍스트 분류 모델의 종류

② 결정 트리(Decision Tree)

- » 데이터의 특징을 이용하여 트리(Tree) 형태의 분류 모델을 만드는 알고리즘
- » 텍스트의 특징을 기준으로 데이터를 분할하여 하위 트리를 생성하는 방식으로 최종적으로 분류를 수행
- » 결정 트리 방식은 인간의 직관적 의사결정 방식과 유사해, 알고리즘의 해석력이 높음

2. 텍스트 분류 모델의 종류

③ 서포트 벡터 머신(Support Vector Machine, SVM)

- » 서포트 벡터 머신 알고리즘은 데이터를 분류하는 결정 경계(decision boundary)를 찾는 알고리즘
- » 분류를 위한 최적의 결정 경계를 찾는 것이 목표
- » 데이터를 고차원 공간으로 매핑(mapping)한 후, 이 공간에서 선형 또는 비선형 결정 경계를 찾아 분류

2. 텍스트 분류 모델의 종류

④ 신경망(Neural Network)

- » 입력층과 출력층 사이에 은닉층을 넣은 신경망으로 분류 모델을 구축
- » 텍스트 데이터를 숫자로 표현할 수 있는 벡터로 변환한다면, 신경망 알고리즘에 적용 가능
- » Bow(Bag-of-words) 방식으로 서로 길이가 다른 텍스트를 고정된 길이의 벡터로 변환한 뒤, 신경망에 텍스트를 입력

3. 베이지안 필터 구현

① 결합 확률

- » 두 개 이상의 사건이 동시에 일어날 확률
- » 두개 이상의 확률 변수를 가짐

ex

주사위 2개(A, B)를 던질 때의 확률

$$P(A, B)$$

3. 베이지안 필터 구현

① 결합 확률

ex

주사위 2개(A, B)를 던졌을 때 A는 5, B는 3가 나올 확률

$$P(A=5, B=3)$$

- » 위의 경우처럼 각각의 사건이 서로에게 영향을 끼치지 않는 경우 확률에서는 이를 “독립” 이라고 부르고 아래의 조건을 만족

$$P(A, B) = P(A)P(B)$$

3. 베이지안 필터 구현

② 조건부 확률

» 어떤 확률이 발생했을 때, 다른 어떤 한 사건이 발생할 확률

$$P(x|y) = \frac{P(x, y)}{P(y)}, P(x, y) = P(x|y)P(y)$$

ex

- 비가 내릴 확률: $P(\text{비})$
- 교통 사고가 발생할 확률: $P(\text{교통사고})$
- 비가 내리는 날에 교통 사고가 발생할 확률: $P(\text{교통사고}|\text{비})$

3. 베이지안 필터 구현

③ 베이즈 정리(Bayes' theorem)

- » 조건부 확률(conditional probability)을 이용하여 사건의 확률을 계산하는 방법
- » 어떤 사건이 일어났을 때,
그 사건이 일어나기 전의 확률을 역추적하여 계산하는 것을 가능하게 함
- » $P(y | x)$ 를 아는 상태에서 $P(x | y)$ 를 구해야 하는 상태에서 쓰임

$$\underbrace{P(x|y)}_{\text{사후 확률(posterior)}} = \frac{\underbrace{P(x)}_{\text{사전 확률(prior)}} P(y|x)}{P(y)}$$

3. 베이esian 필터 구현

③ 베이즈 정리(Bayes' theorem)

📦 베이즈 정리를 이용한 사후확률 계산의 예

밸런타인데이에 상대방에게 초콜릿을 받았다면,
상대방이 나를 좋아할 확률은?

- $P(\heartsuit) =$ 좋아하지 않을 확률
- $P(\blackheartsuit) =$ 좋아할 확률
- $P(\square) =$ 초콜릿을 주지 않을 확률
- $P(\blacksquare) =$ 초콜릿을 줄 확률

3. 베이esian 필터 구현

③ 베이즈 정리(Bayes' theorem)

📦 베이즈 정리를 이용한 사후확률 계산의 예

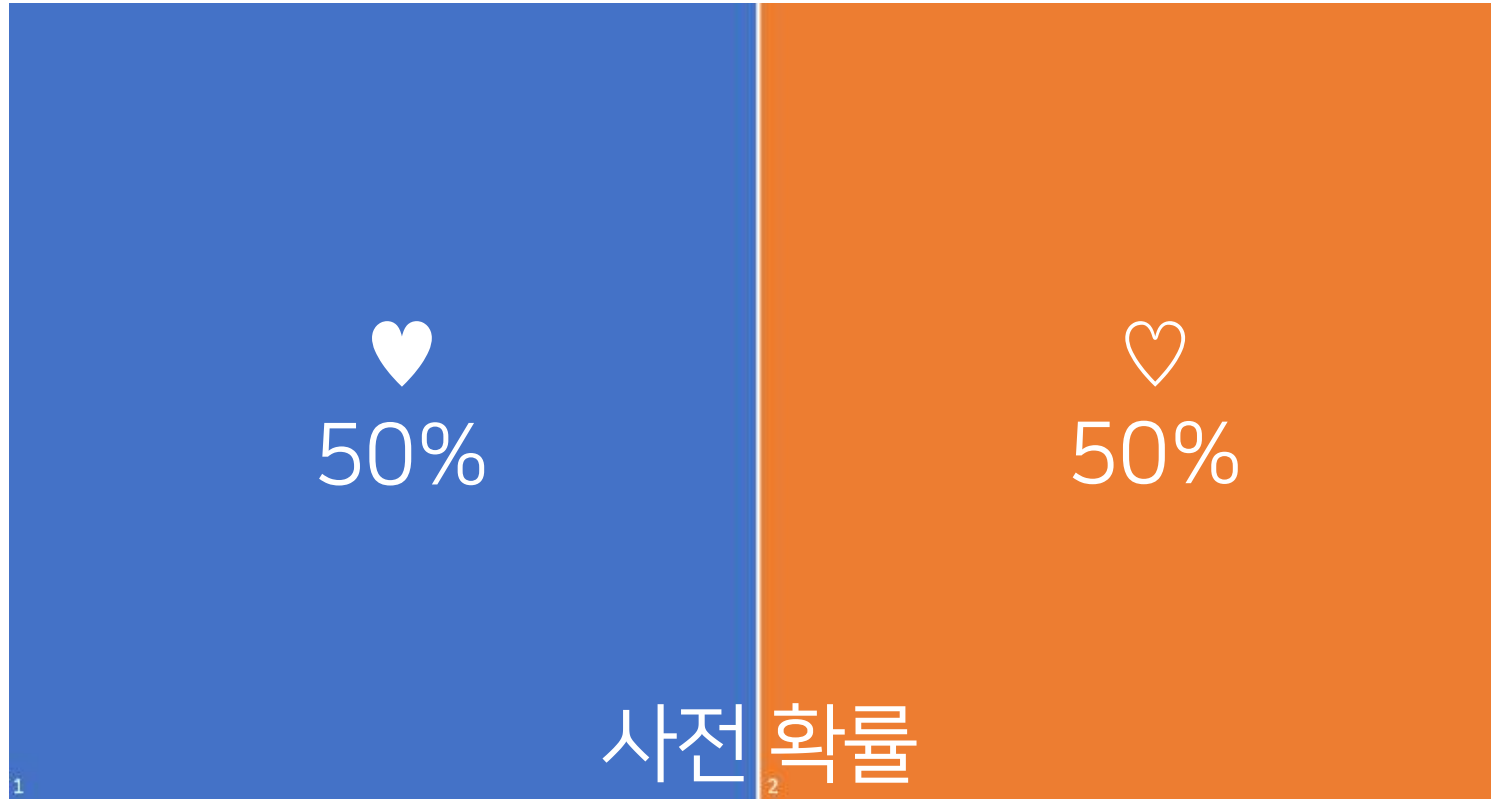
밸런타인데이에 상대방에게 초콜릿을 받았다면,
상대방이 나를 좋아할 확률은?

$$P(\heartsuit|\blacksquare) = \frac{P(\heartsuit)P(\blacksquare|\heartsuit)}{P(\blacksquare)}$$

3. 베이지안 필터 구현

③ 베이즈 정리(Bayes' theorem)

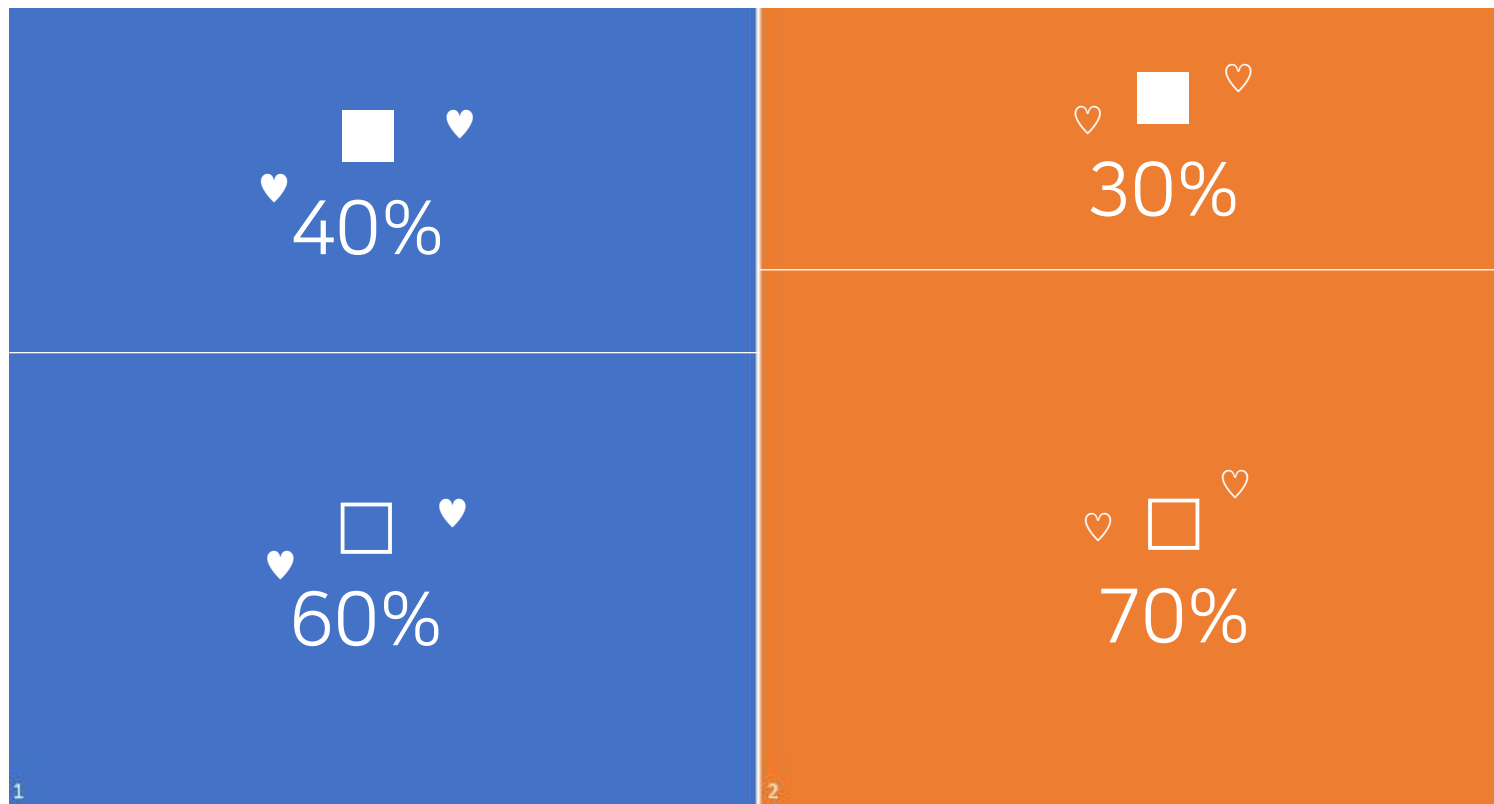
④ 베이즈 정리를 이용한 사후확률 계산의 예



3. 베이지안 필터 구현

③ 베이즈 정리(Bayes' theorem)

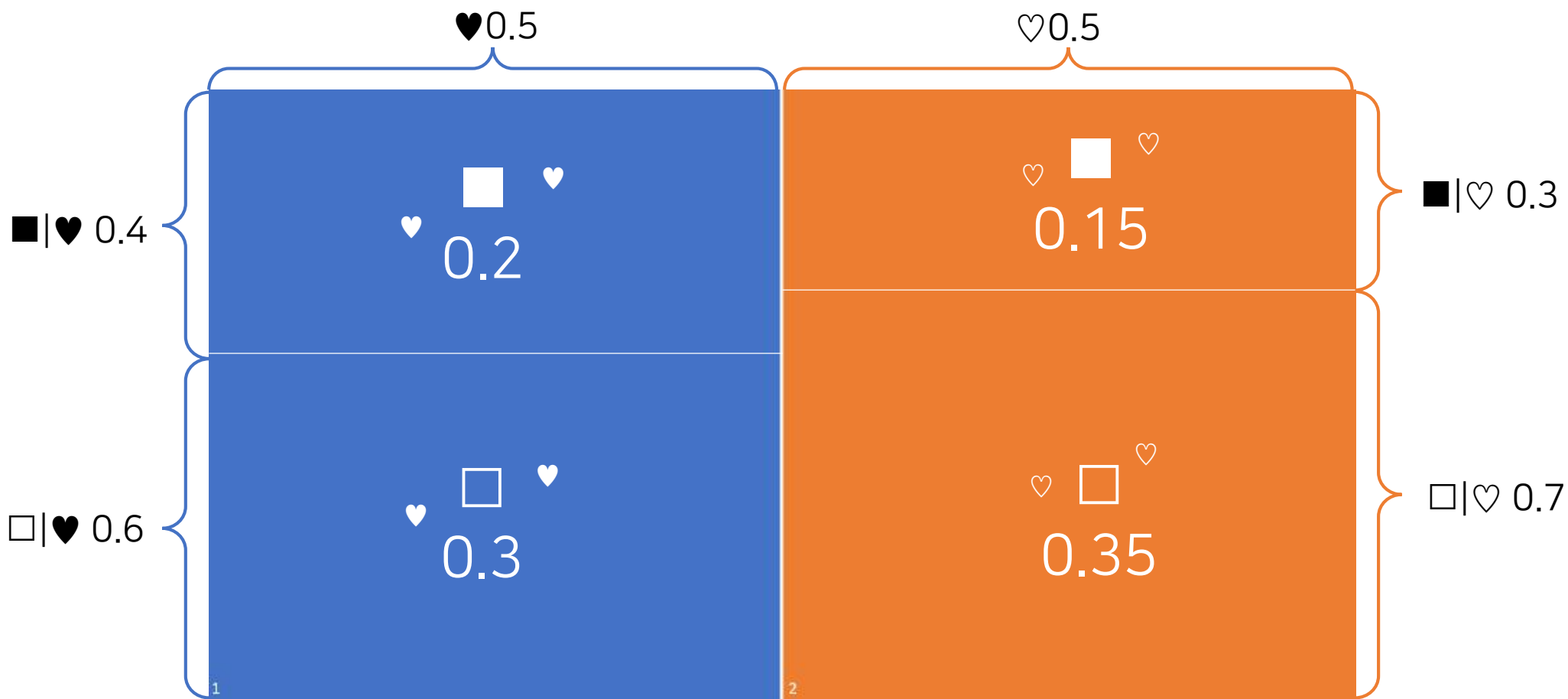
④ 베이즈 정리를 이용한 사후확률 계산의 예



3. 베이지안 필터 구현

③ 베이즈 정리(Bayes' theorem)

📦 베이즈 정리를 이용한 사후확률 계산의 예



3. 베이지안 필터 구현

③ 베이즈 정리(Bayes' theorem)

④ 베이즈 정리를 이용한 사후확률 계산의 예

$$P(\heartsuit|\blacksquare) = \frac{P(\heartsuit)P(\blacksquare|\heartsuit)}{P(\blacksquare)} = \frac{20}{35} \approx 57\%$$

3. 베이지안 필터 구현

④ 나이브 베이즈 분류(Naïve Bayes Classification)

- » 확률적 기계학습 알고리즘으로, 베이즈 정리에 기반하여 데이터를 분류하는 알고리즘
- » 특성(feature)을 여러 개 사용하여 판단 근거를 더 높이는 방법으로 데이터를 분류함

$$\hat{y} = \operatorname{argmax}_{k \in \{1, 2, \dots, k\}} P(c_k) \prod_{i=1}^n P(x_i | c_k)$$

3. 베이지안 필터 구현

④ 나이브 베이즈 분류(Naïve Bayes Classification)

📦 나이브 베이즈 분류를 이용한 분류의 예시

$$P(\heartsuit | \text{초콜릿} \cap \text{생일선물} \cap \text{먼저 연락}) = P(\heartsuit)P(\text{초콜릿} | \heartsuit)P(\text{생일선물} | \heartsuit)P(\text{먼저 연락} | \heartsuit)$$

∨

$$P(\heartsuit | \text{초콜릿} \cap \text{생일선물} \cap \text{먼저 연락}) = P(\heartsuit)P(\text{초콜릿} | \heartsuit)P(\text{생일선물} | \heartsuit)P(\text{먼저 연락} | \heartsuit)$$

≈



3. 베이지안 필터 구현

④ 나이브 베이즈 분류(Naïve Bayes Classification)

📦 나이브 베이즈 분류를 이용한 텍스트 분류

- 텍스트의 단어를 토큰화 하여 이 단어들을 나이브 베이즈 분류기로 입력
- 나이브 베이즈 분류기는 모든 단어가 독립적이라고 가정

$$P(\text{정상 메일} \mid \text{입력 텍스트}) = P(t1 \mid \text{정상 메일}) \times P(t2 \mid \text{정상 메일}) \times P(t3 \mid \text{정상 메일}) \times P(\text{정상 메일})$$

$$P(\text{스팸 메일} \mid \text{입력 텍스트}) = P(t1 \mid \text{스팸 메일}) \times P(t2 \mid \text{스팸 메일}) \times P(t3 \mid \text{스팸 메일}) \times P(\text{스팸 메일})$$

3. 베이지안 필터 구현

④ 나이브 베이즈 분류(Naïve Bayes Classification)

📦 나이브 베이즈 분류를 이용한 스팸 분류기의 예

입력 예

you free lottery

Training Data

번호	메일로부터 토큰화 및 정제 된 된 단어들	분류
1	me free lottery	스팸 메일
2	free get free you	스팸 메일
3	you free scholarship	정상 메일
4	free to contact me	정상 메일
5	you won award	정상 메일
6	you ticket lottery	스팸 메일

3. 베이지안 필터 구현

④ 나이브 베이즈 분류(Naïve Bayes Classification)

📦 나이브 베이즈 분류를 이용한 스팸 분류기의 예

$$\begin{aligned} P(\text{정상 메일} \mid \text{입력 텍스트}) &= P(\text{you} \mid \text{정상 메일}) \times P(\text{free} \mid \text{정상 메일}) \times P(\text{lottery} \mid \text{정상 메일}) \times P(\text{정상 메일}) \\ P(\text{스팸 메일} \mid \text{입력 텍스트}) &= P(\text{you} \mid \text{스팸 메일}) \times P(\text{free} \mid \text{스팸 메일}) \times P(\text{lottery} \mid \text{스팸 메일}) \times P(\text{스팸 메일}) \end{aligned}$$

≈

$$\begin{aligned} P(\text{정상 메일} \mid \text{입력 텍스트}) &= P(\text{you} \mid \text{정상 메일}) \times P(\text{free} \mid \text{정상 메일}) \times P(\text{lottery} \mid \text{정상 메일}) \\ P(\text{스팸 메일} \mid \text{입력 텍스트}) &= P(\text{you} \mid \text{스팸 메일}) \times P(\text{free} \mid \text{스팸 메일}) \times P(\text{lottery} \mid \text{스팸 메일}) \end{aligned}$$

||

$$\begin{aligned} P(\text{정상 메일} \mid \text{입력 텍스트}) &= 2/10 \times 2/10 \times 0/10 = 0 \\ P(\text{스팸 메일} \mid \text{입력 텍스트}) &= 2/10 \times 3/10 \times 2/10 = 0.012 \end{aligned}$$



개발 실무 실습하기

P r a c t i c a l P r o g r a m m i n g f o r A I

텍스트 분류 실습

① 텍스트 분류란?

주어진 텍스트를 미리 정의된
카테고리 중 하나로 자동 분류하는 과정

② 텍스트 분류 모델의 종류

나이브 베이즈

결정 트리

서포트 벡터 머신

신경망

③ 베이지안 필터 구현

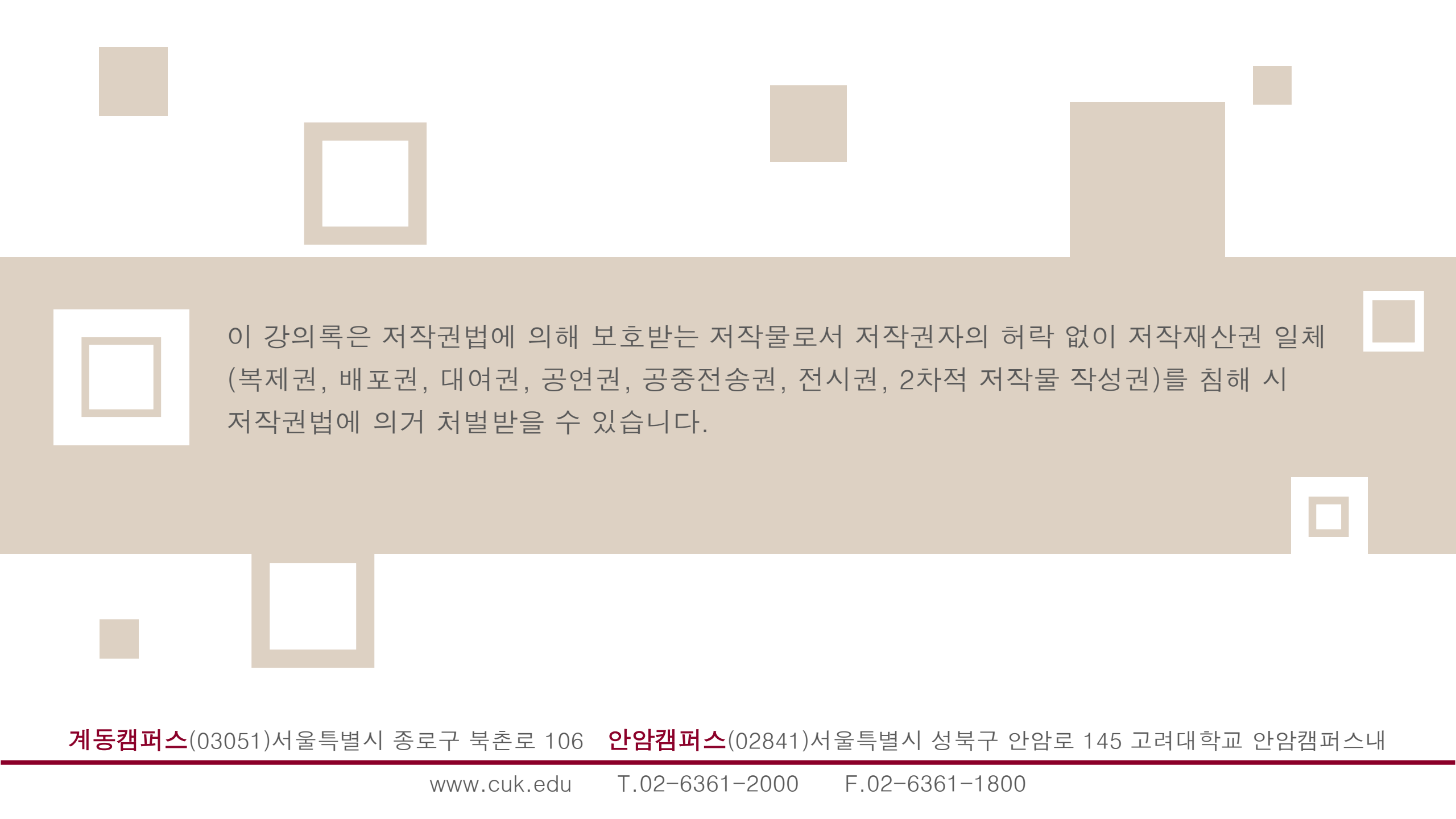
베이즈 정리 기반의 나이브베이즈 분류기 활용

④ 텍스트 분류 실습

베이지안 필터를 이용한 텍스트 분류

참고문헌

- ❏ 딥러닝을 이용한 자연어 처리 입문(<https://wikidocs.net/book/2155>)
- ❏ <https://www.youtube.com/watch?v=KueAHu7iFNE>



이 강의록은 저작권법에 의해 보호받는 저작물로서 저작권자의 허락 없이 저작재산권 일체 (복제권, 배포권, 대여권, 공연권, 공중전송권, 전시권, 2차적 저작물 작성권)를 침해 시 저작권법에 의거 처벌받을 수 있습니다.