

# 주요 포트폴리오 설명

---

---

중앙대학교 대학원

---

전자상거래 및 인터넷 응용 연구실

---

김 준 호

---







## Mostly solved

### Spam detection

Let's go to Agra!



Buy V1AGRA ...



### Part-of-speech (POS) tagging

ADJ ADJ NOUN VERB ADV

Colorless green ideas sleep furiously.

### Named entity recognition (NER)

PERSON ORG LOC

Einstein met with UN officials in Princeton

## Making good progress

### Sentiment analysis

Best roast chicken in San Francisco!



The waiter ignored us for 20 minutes.



### Coreference resolution

Carter told Mubarak he shouldn't run again.

### Word sense disambiguation (WSD)

I need new batteries for my *mouse*.



### Parsing

I can see Alcatraz from the window!

### Machine translation (MT)

第13届上海国际电影节开幕...



The 13<sup>th</sup> Shanghai International Film Festival...

### Information extraction (IE)

You're invited to our dinner party, Friday May 27 at 8:30



Party  
May 27  
add

## Still really hard

### Question answering (QA)

Q. How effective is ibuprofen in reducing fever in patients with acute febrile illness?

### Paraphrase

XYZ acquired ABC yesterday

ABC has been taken over by XYZ

### Summarization

The Dow Jones is up

The S&P500 jumped

Housing prices rose



Economy is good

### Dialog

Where is Citizen Kane playing in SF?



Castro Theatre at 7:30. Do you want a ticket?





## Mostly solved

Spam detection

Let's go to the beach. **스팸 분류** ✓

Buy V1AGRA ... ✗

Part-of-speech (POS) tagging

ADJ 품사 결정 ADV

Colorless green ideas sleep furiously.

Named entity recognition (NER)

PERSON 개체명 인식 LOCATION

Einstein met with UN officials in Princeton

## Making good progress

Sentiment analysis

Best road to the beach. **감성 분석** 👍

The waiter ignored us for 20 minutes. 👎

Coreference resolution

Carter told Mubarak he shouldn't run again.

Word sense disambiguation (WSD)

I need new batteries for my **mouse**.

Parsing

I can see Alcatraz from the window! **구문 분석**

Machine translation (MT)

第13届上海国际电影节开幕

The 13<sup>th</sup> Shanghai International Film Festival... **기계 번역** ➡

Information extraction (IE)

You're invited to a film party, Friday May 27 at 8:30 **정보 추출**

Party May 27 add

## Still really hard

Question answering (QA)

Q. How effective is paracetamol in reducing fever in patients with acute febrile illness? **질의 응답**

Paraphrase

XYZ acquired ABC yesterday. **의역**

ABC has been taken over by XYZ

Summarization

The Dow Jones is up. **요약**

The S&P500 jumped

Housing prices rose

Economy is good

Dialog

Where is Citizen Kane playing in SF?

Castro Theatre at 7:30. Do you want a ticket? **대화**



## Mostly solved

Spam detection

Let's go to Agra! ✓

Buy V1AGRA ... ✗

Part-of-speech (POS) tagging

ADJ 품사 결정 ADV

Colorless green ideas sleep furiously.

Named entity recognition (NER)

PERSON ORG LOC

Einstein met with UN officials in Princeton

## Making good progress

Sentiment analysis

Best road to S. ✓

The waiter ignored us for 20 minutes. ✗

Coreference resolution

Carter told Mubarak he shouldn't run again.

Word sense disambiguation (WSD)

I need new batteries for my mouse.

Parsing

I can see Alcatraz from the window!

Machine translation (MT)

第13届上海国际电影节开幕...

The 13<sup>th</sup> Shanghai International Film Festival...

Information extraction (IE)

You're invited to a party, Friday May 27 at 8:30

Party May 27 add

## Still really hard

Question answering (QA)

질의 응답

Q. How effective is ... reducing fever in patients with acute febrile illness?

Paraphrase

의역

XYZ acquired ABC yesterday.

ABC has been taken over by XYZ

Summarization

The Dow Jones is up

The S&P500 jumped

Housing prices rose

Economy is good

Dialog

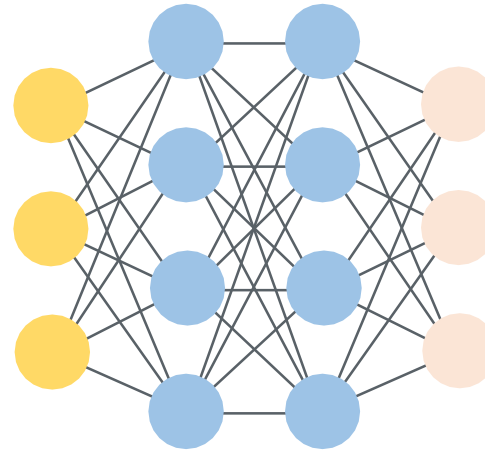
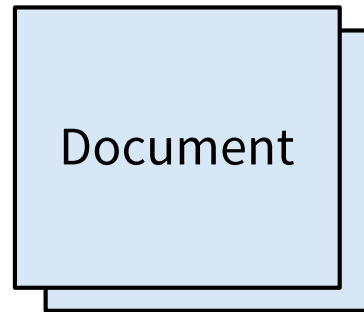
Where is Citizen Kane playing in SF?

Castro Theatre at 7:30. Do you want a ticket?

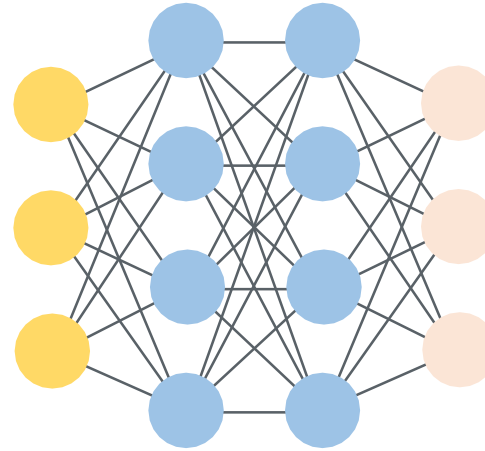
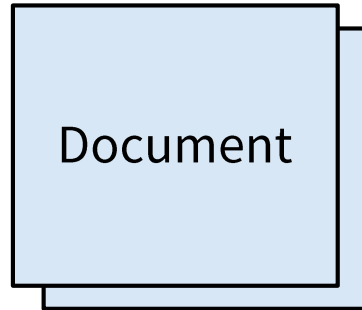
Lotto...



# Skill & Process



Output



Output

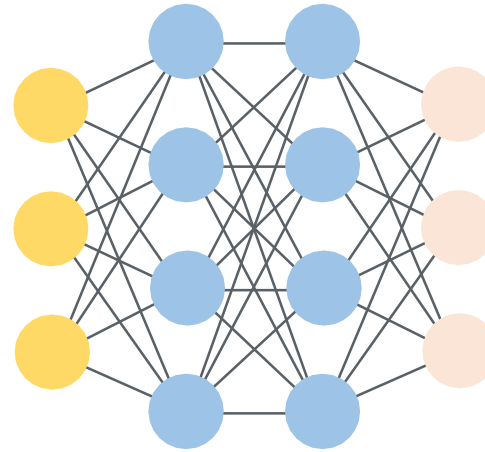
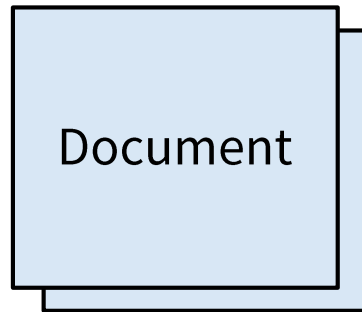
jsoup parser

Hadoop

Hannanum, Komoran, Twitter 형태소분석기

TF-IDF





Output

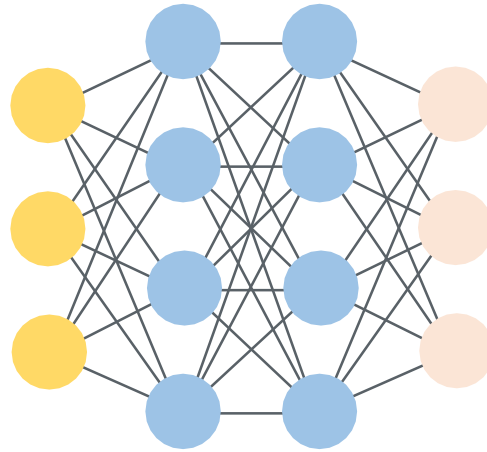
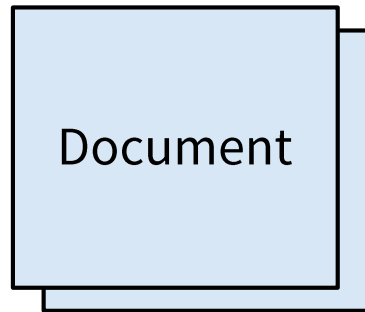
jsoup parser 수집

Hadoop 저장

Hannanum, Komoran, Twitter 형태소분석기

TF-IDF

전처리



Output

jsoup parser 수집

Hadoop 저장

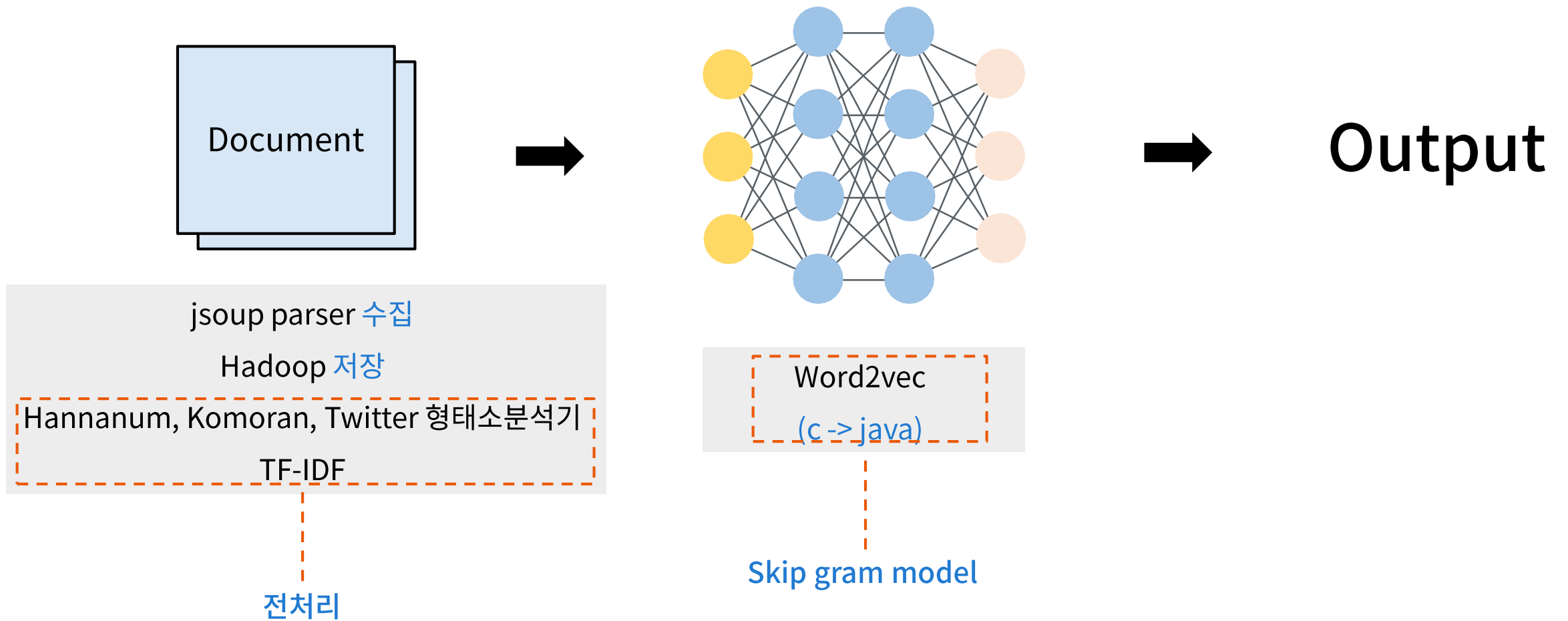
Hannanum, Komoran, Twitter 형태소분석기

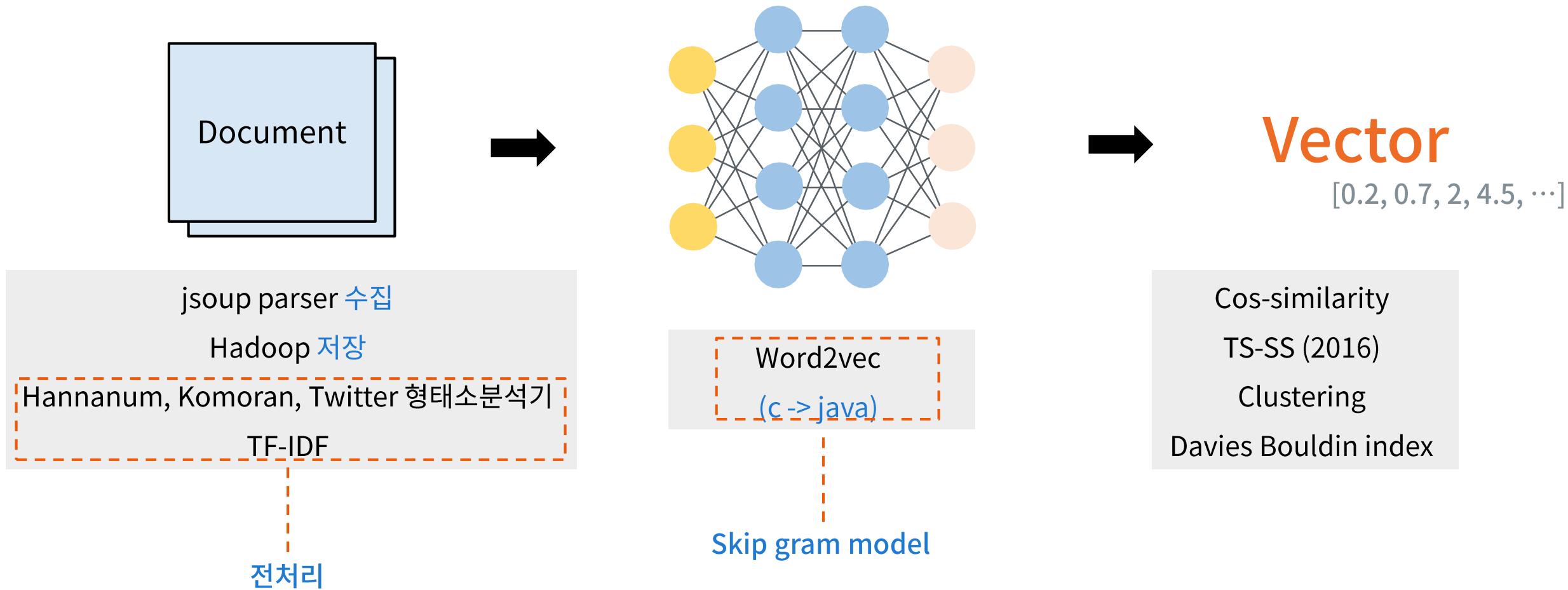
TF-IDF

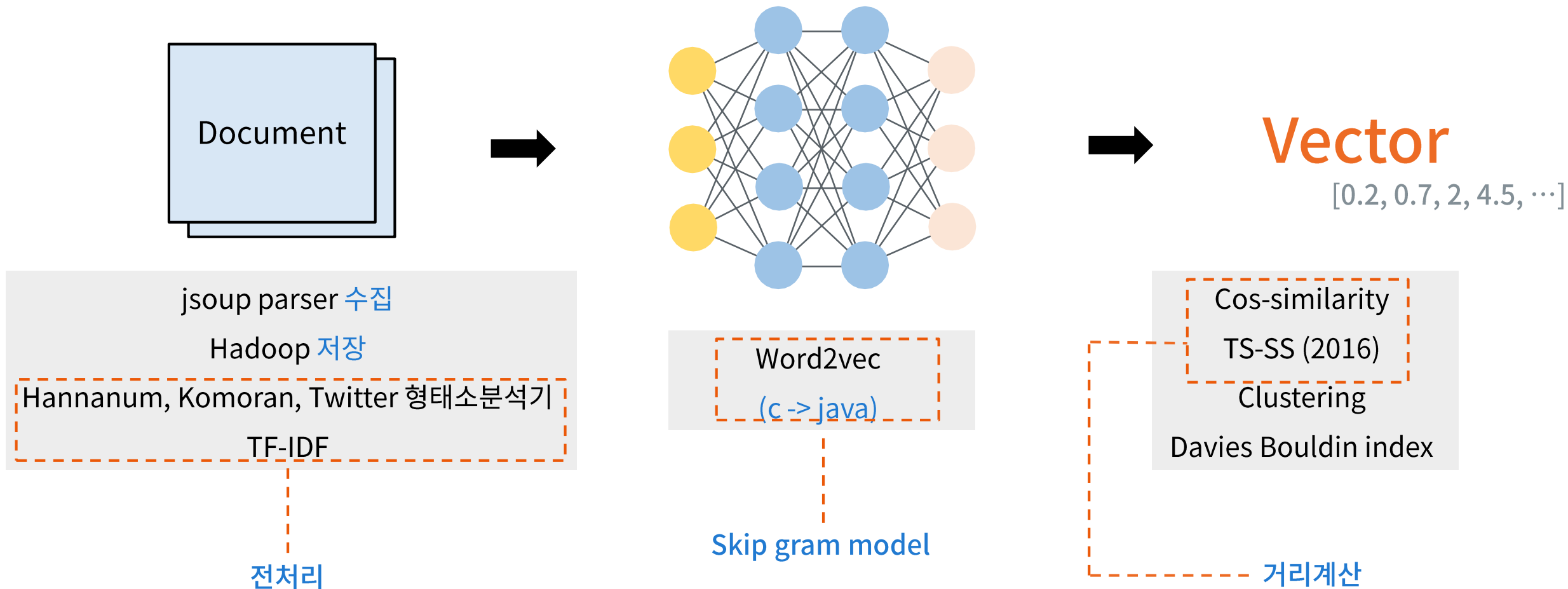
전처리

Word2vec

(c -> java)









# Project / 품사 결정, 언어 중의성

Project / 품사 결정, 언어 중의성

연구원들이

연구를

효율적으로 하게 하자 !

## Step 1 : 컨셉

아이폰





## Step 1 : 컨셉



## Step 1 : 컨셉

아이폰



브랜드

애플, 안드로이드, 구글 ...



관련기술

터치스크린, 인벤터 ...



# How ?

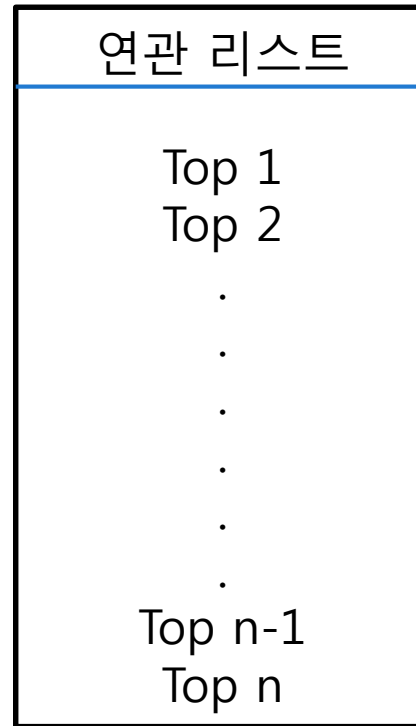
단어

0.2, 3, 0.5, ...

Word2Vec

0.2, 3, 0.5, ...

Word2Vec



cosine similarity

0.2, 3, 0.5, ...

Word2Vec



연관 리스트

Top 1

Top 2

.

.

.

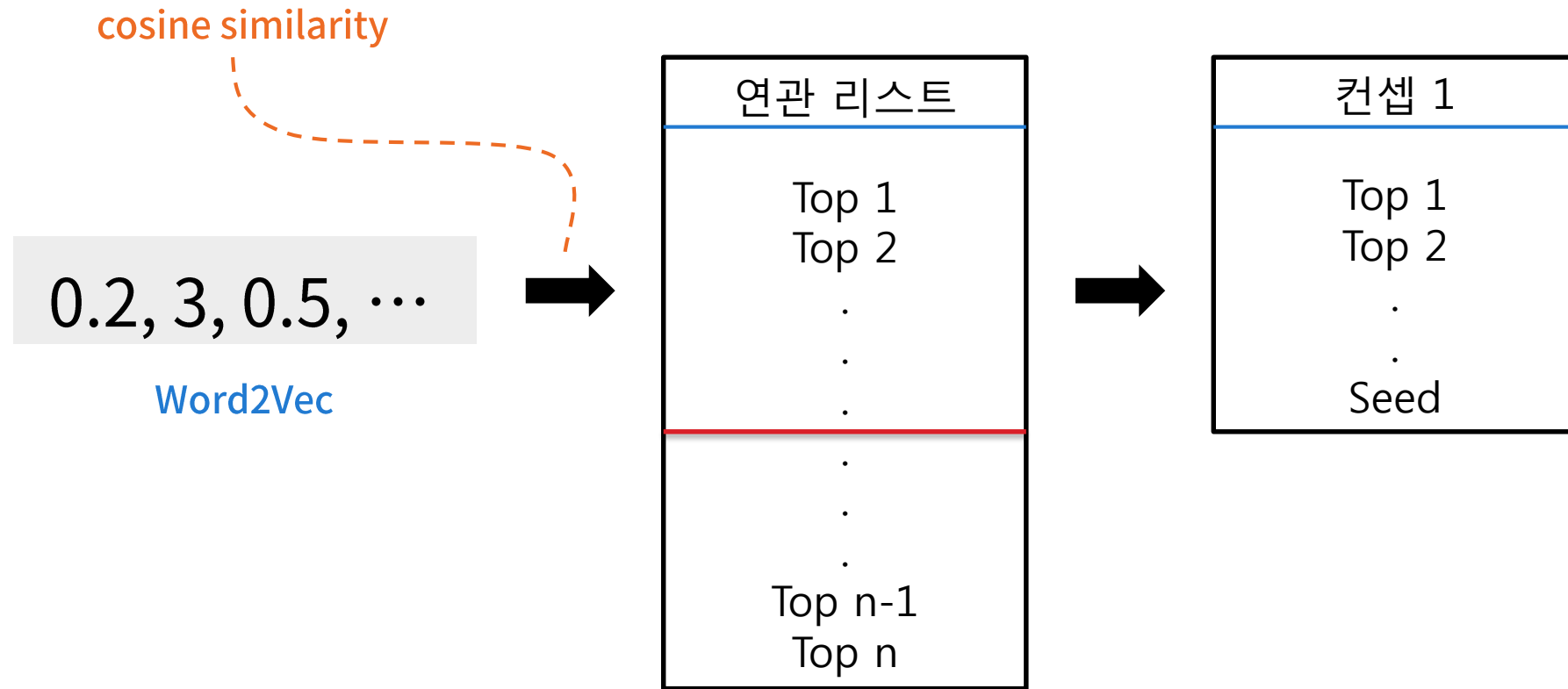
.

.

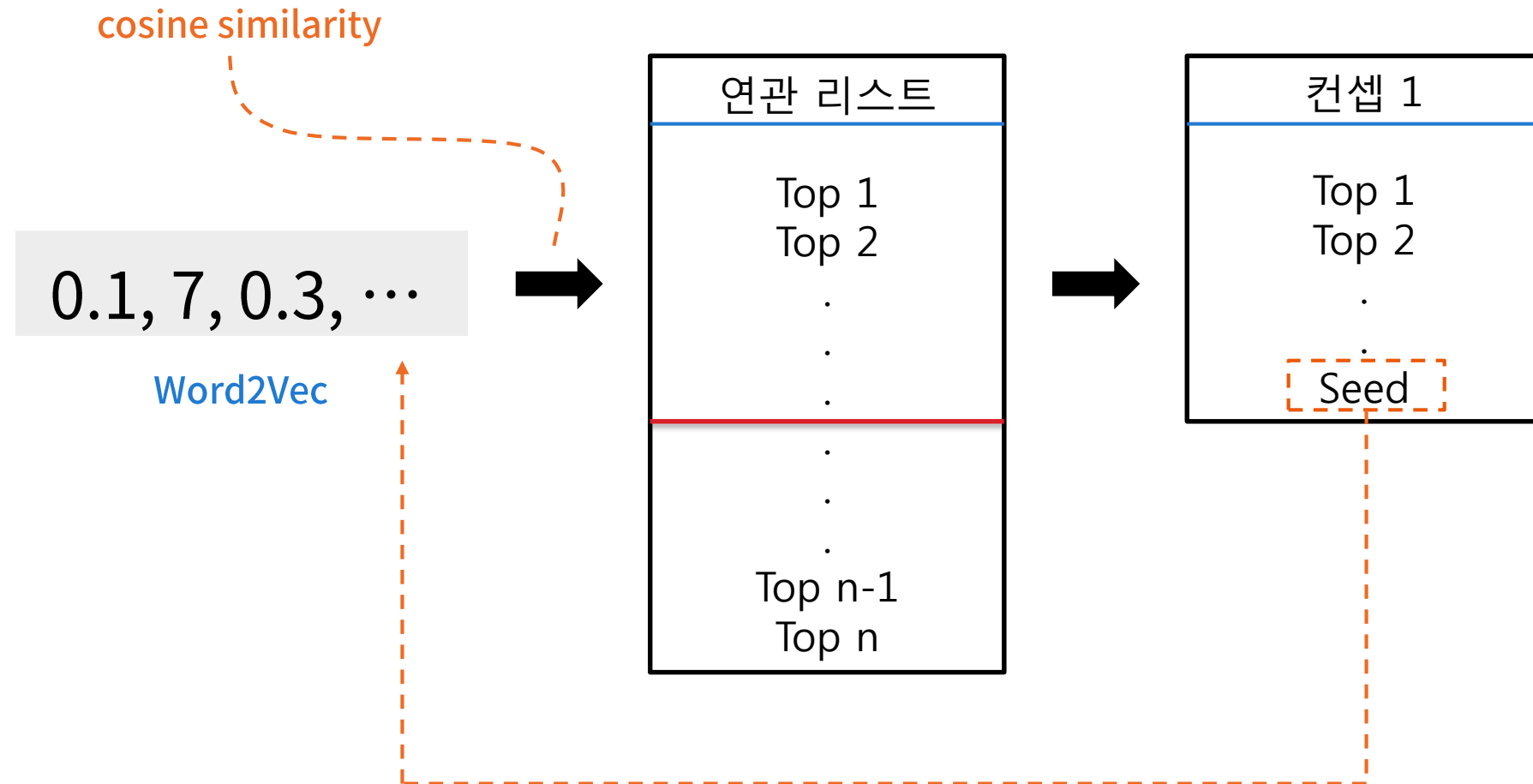
.

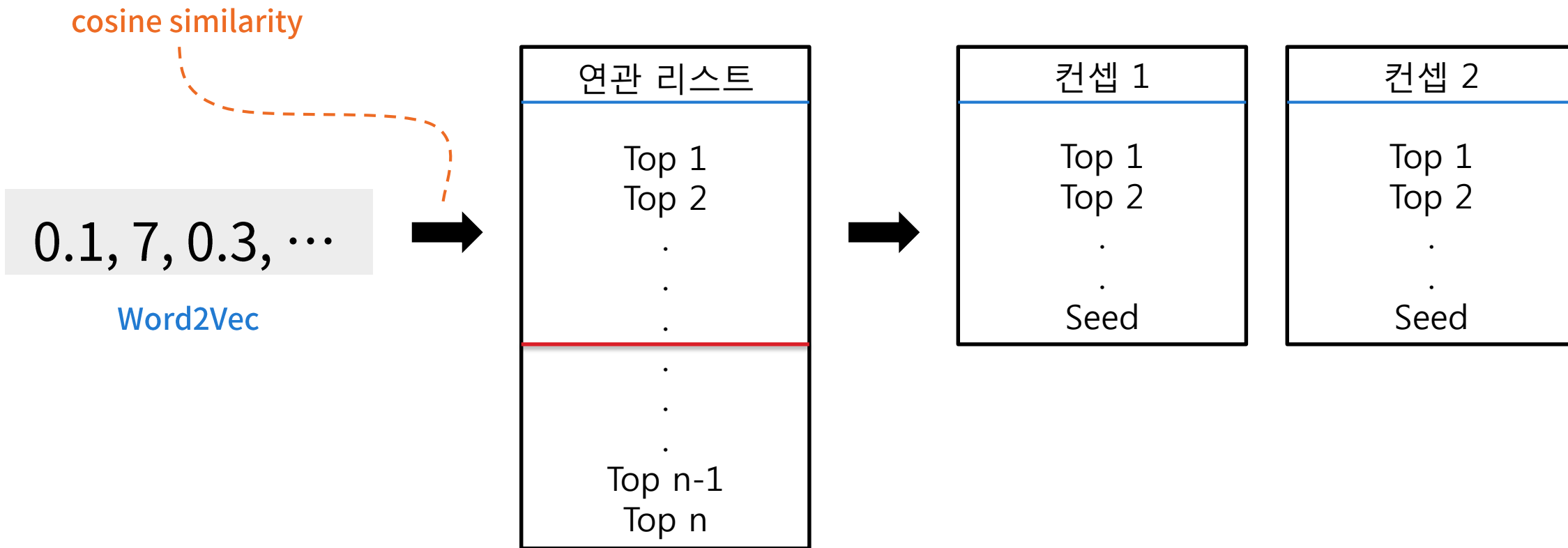
Top n-1

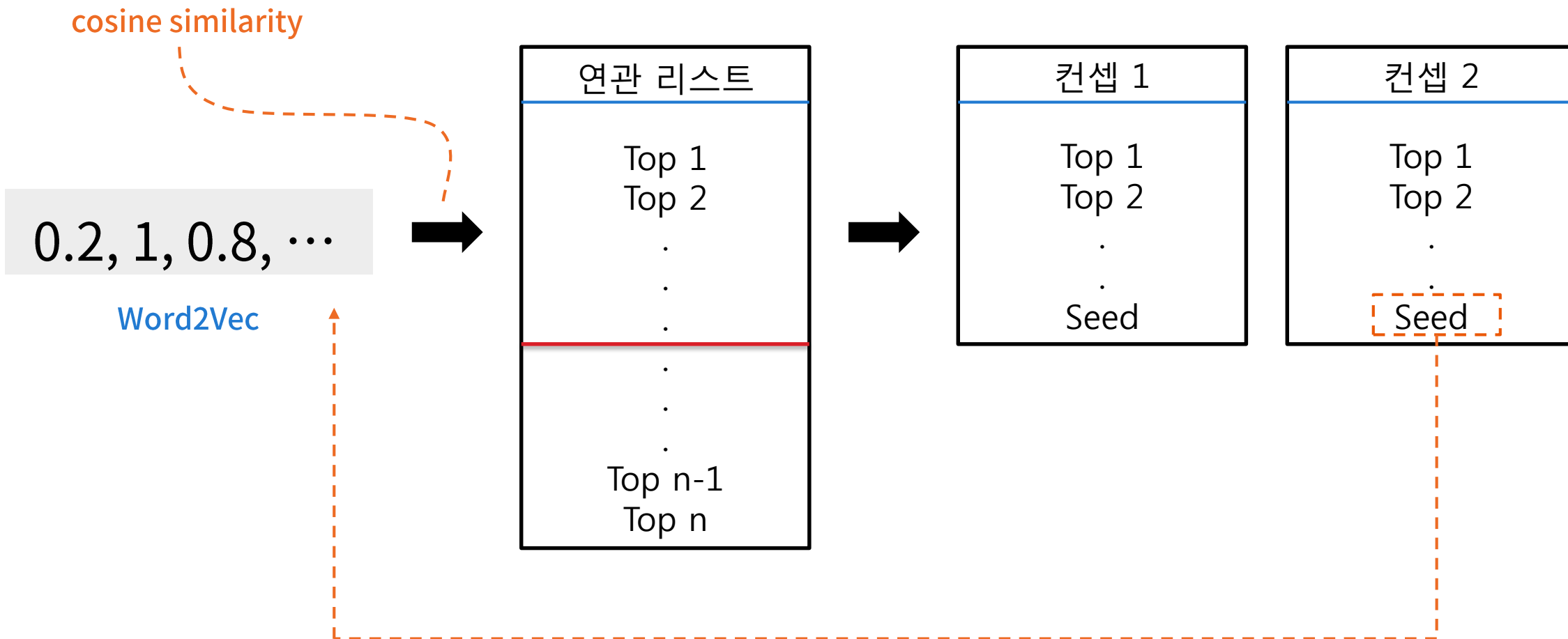
Top n











# Infinity...

$$\textit{Cluster score} = \frac{1}{n} \sum_{i=1}^n \max_{j \neq i} \left( \frac{\sigma_i + \sigma_j}{d(c_i, c_j)} \right)$$

$$Cluster\ score = \frac{1}{n} \sum_{i=1}^n \max_{j \neq i} \left( \frac{\sigma_i + \sigma_j}{d(c_i, c_j)} \right)$$

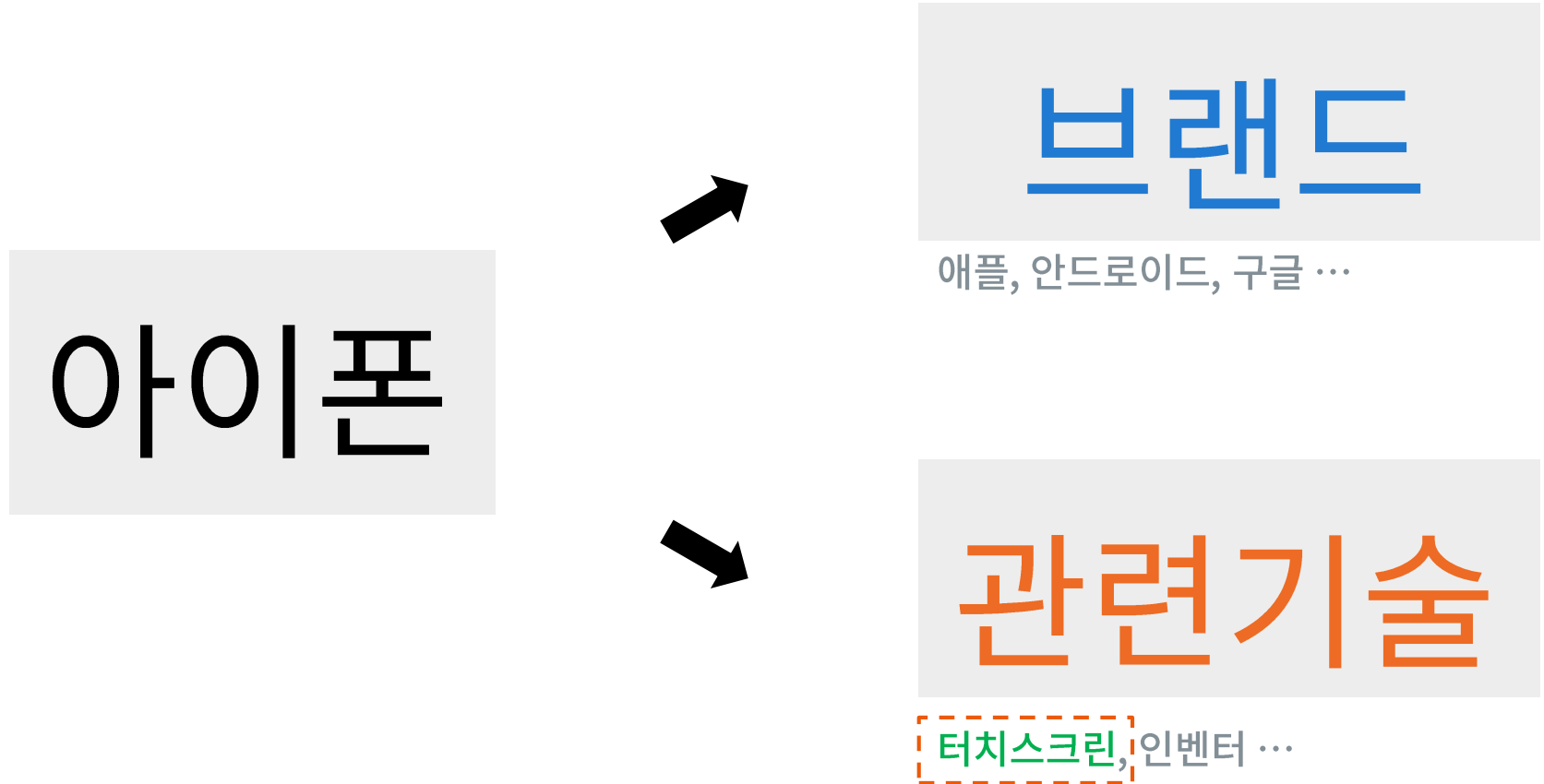
클러스터 내 유사도

클러스터 간 유사도

## Step 2 : 시계열 분석



## Step 2 : 시계열 분석





## Step 2 : 시계열 분석

터치스크린

## Step 2 : 시계열 분석

터치스크린



2009-2010	2011-2012	2013-2014
패널	패널	패널
용량식	용량식	
저항막	저항막	
정전식	정전식	
		온셀
		인셀

## Step 2 : 시계열 분석

터치스크린

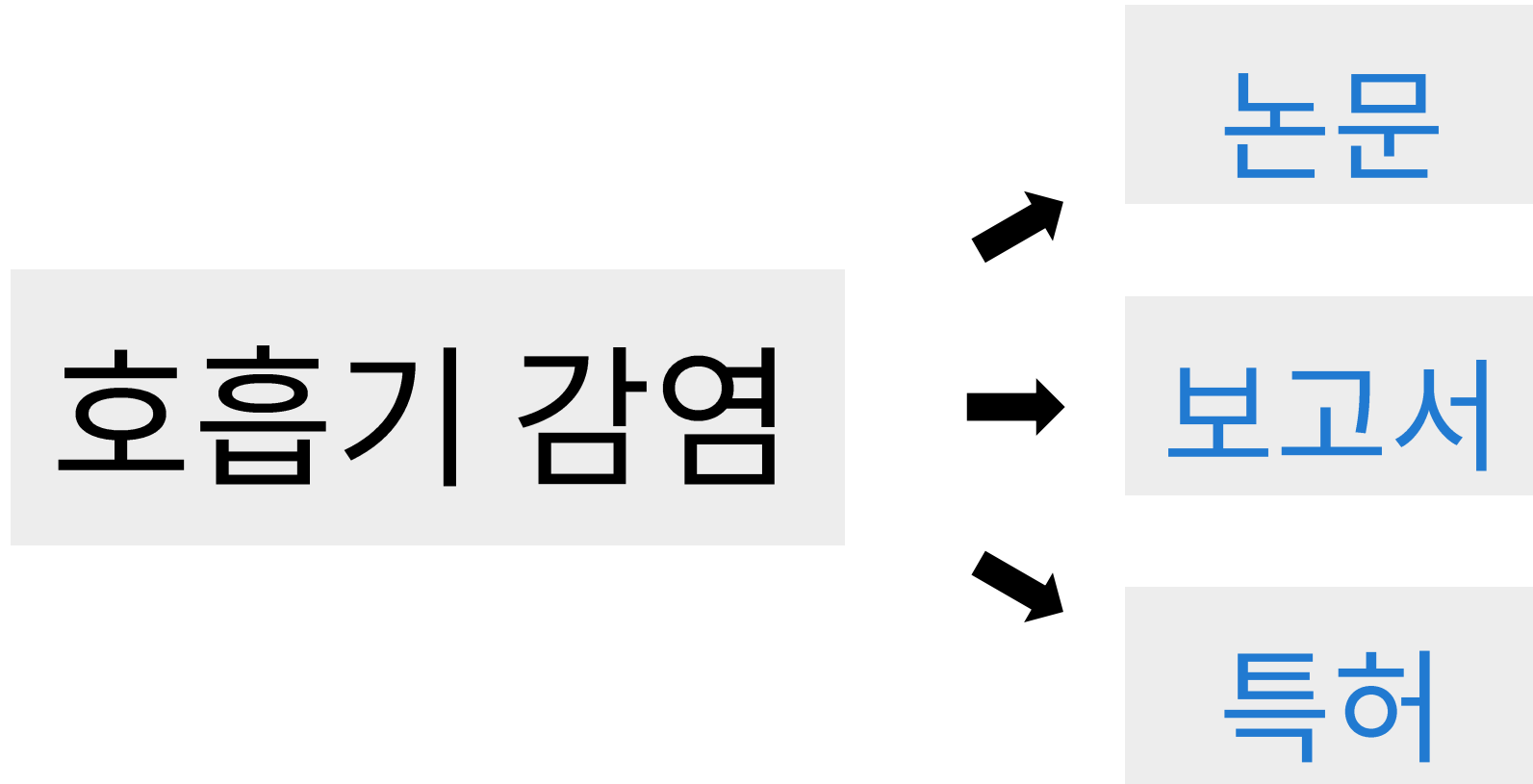


2009-2010	2011-2012	2013-2014
패널	패널	패널
용량식	용량식	
저항막	저항막	
정전식	정전식	
		온셀
		인셀

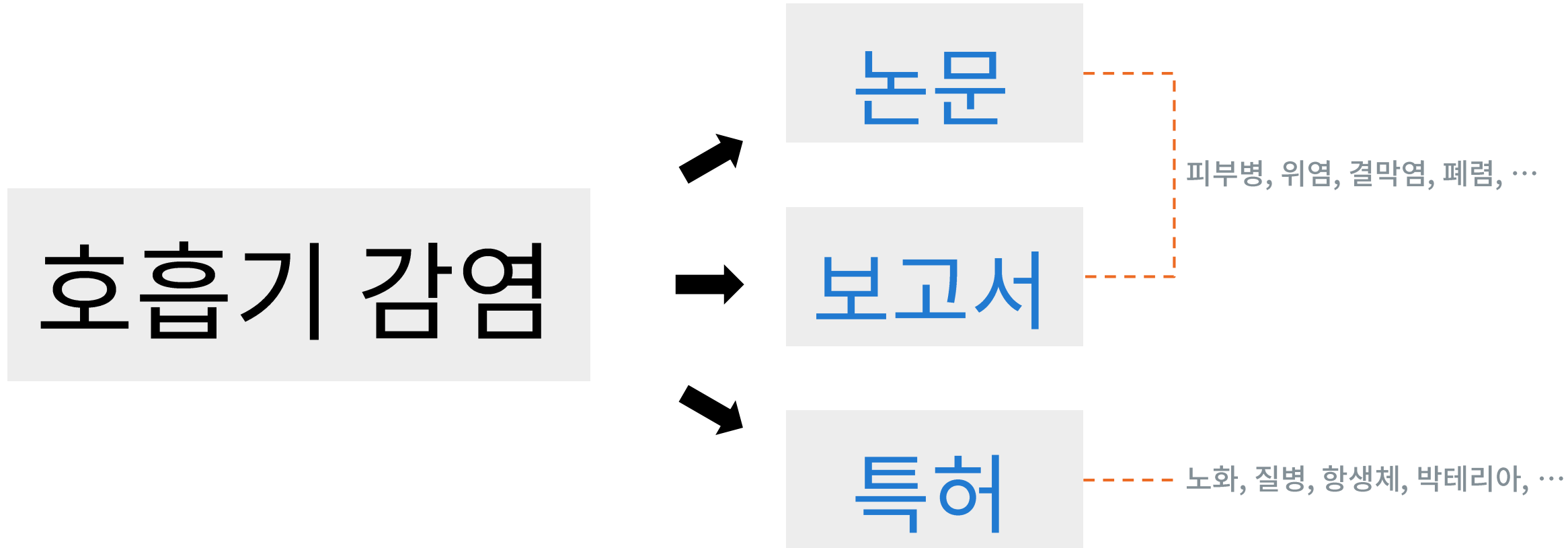
## Step 3 : 분야별 분석

호흡기 감염

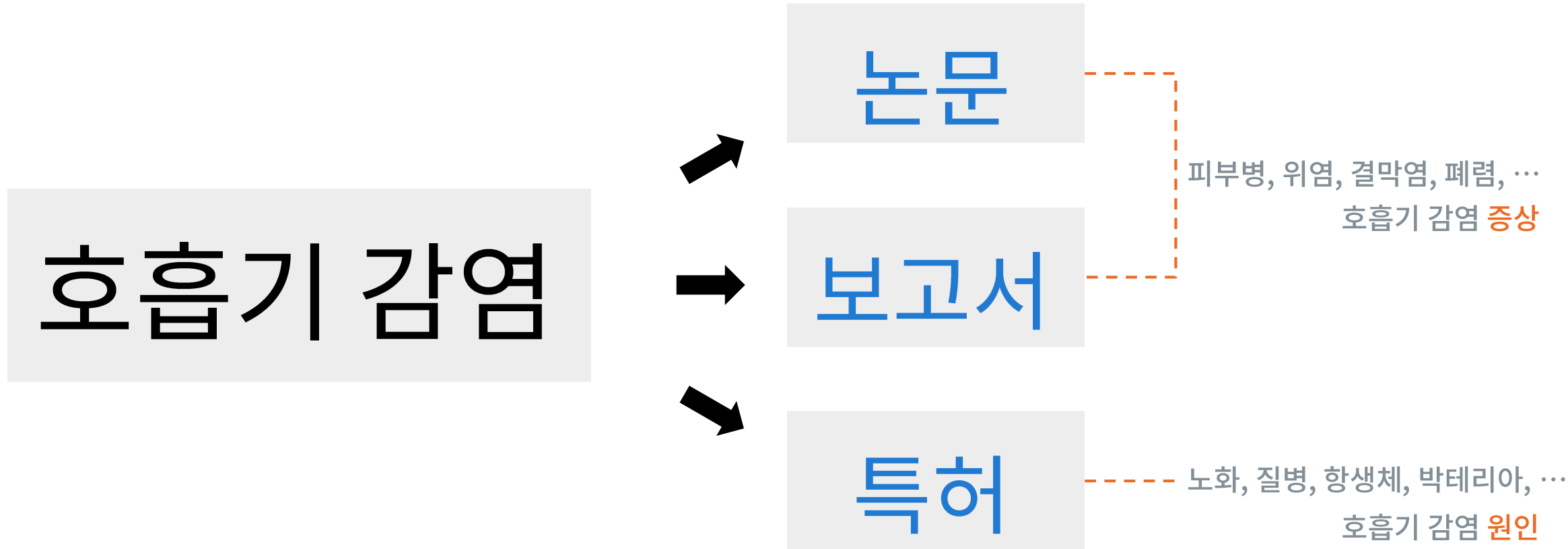
## Step 3 : 분야별 분석



## Step 3 : 분야별 분석



## Step 3 : 분야별 분석





# 서비스 화면





연관 검색 엔진

☒ 보건 ☐ 반도체 ☐ 나노

의사 결정

search

KeyWord	score
다요소	0.5973275750506075
다기준	0.5556915560532243
결정법	0.4923757182550543
가중	0.34897902287843385

KeyWord	score
선택	0.4727749051003636
분석가	0.4627437880585548
합리	0.44997393772941974
정확	0.4474464100280191
경영자	0.44112246867626337
설명	0.43229386140246473
속성	0.4209597289756161
트리	0.41484470441953447
추론	0.41390301647285926
마코프	0.41025127907298464

KeyWord	score
결정권자	0.5646898522807938
결정부	0.5260000315745299
다속	0.41511318112472473
설계자	0.40299522220759443
우선순위	0.3983141403887548
집합	0.3901025307339995
역행렬	0.3879186328180932
객관	0.38776494015259083
상수관망	0.3840038911569563
주주	0.3822714047407798

논문

특허

연구보고서

분석

특허 검색 결과

KOR1020110114576	의사 결정 추천 시스템 및 방법 <a href="#">[초록]</a>	성균관대학교
KOR1020040005141	의사 결정 및 감시 기능을 가지는 지능형 에이전트를구비한 무선통신 단말기 및 그 방법 <a href="#">[초록]</a>	주식회사 팬택
KOR1020010036579	의사 결정 및 전략 회의 장치 및 방법 <a href="#">[초록]</a>	오를리스 주식
KOR1020130155896	임상 의사 결정 지원 방법 및 서버 장치 <a href="#">[초록]</a>	에스케이텔레콤
KOR1020120148207	다중 의사 결정 문제 해결을 위한 구간 회색수 및 엔트로피 기반 해법 연산 방법 <a href="#">[초록]</a>	성균관대학교
KOR1020117025698	프로세스 의사 결정 지원 시스템의 수립 방법 <a href="#">[초록]</a>	지이 인프라스 리카 (프로프리

Research / 감성분석, 질의응답, 의역

식사 논문  
그리고...

## 지식iN 1-8 / 8건



Q [페이코 오프라인 결제 방법](#) 2017.01.07.

편의점에서 **페이코**를 사용하고 싶을때!!! 오프라인에서 **페이코**를 사용할때는... Bc카드를 **페이코**를 통해 편의점에서 사용법을 자세하게 알려주세요!! 참고로 기정은...

A **NHN**엔터테인먼트 PAYCO dhk CU,,양사는 제휴를 기념해 1월 1일부터 한달간 도시락 및 일부 스낵을 **페이코**로 결제할 경우 500원을 할인해주는 쿠폰을 제공하며 최대 2...

파워지식iN 답변 pleasure | 쇼핑 > 쇼핑물, 시장 | 1



Q [위비카드요](#) 2017.01.02.

포인트형으로 만들고 싶습니다.

A **NHN** Entertainment 서비스 - 우리카드-백스 제휴이용권 50% 청구할인 - 한국인 출전 **결제** 시 20% 청구할인 - 티켓링크 10% 청구할인 - 코미코 20% 청구할인 - **페이코** 포인트 1...

경제 > 금융 > 신용카드 | 5

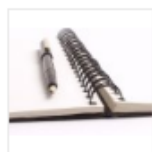


Q [사기당한거 같아요 도와주세요](#) 2016.07.11.

회사는 **NHN**엔터테인먼트, 주식회사 유비핀, playn, 갤럭시마커, 주식회사 마이엠... 보아하니 **페이코**나 이런 **결제**시스템 회사인데 왜 밑도 끝도 없이 이런 큰 돈을...

A 이해가아질않습니다 본인도 모르게 출금이되어다면 회수 **방법**이있습니다 자세한 사항은 쪽지로 연락처 남겨두시면 자세히 상담후 도움드립니다 저희 와이프도 유비핀...

형사사건 > 재산범죄 > 사기 | 2



Q [nhn한국사이버결제 보유중입니다. 어떻게야 할지 속 시원히...](#) 2016.12.06.

**nhn**한국사이버**결제** 보유중입니다. 어떻게야 할지 속 시원히 답변 부탁드립니다. 어찌 하오리까 ??

A 확보중에있구요 **nhn**한국사이버**결제**는 **페이코** 출시 효과에 대한 과도한 시장... 가장 합리적인 주식 투자는 이렇게 올바른 가치 투자 **방법**으로 꾸준히 인내하며 투자를...

경제 > 금융 > 주식, 증권 | 2



NAVER

NHN 페이코 결제방법

검색

통합검색

블로그

카페

지식iN

이미지

동영상

어학사전

뉴스

더보기

지식iN 1-8 / 8건



Q [페이코 오프라인 결제 방법](#) 2017.01.07.

편의점에서 **페이코**를 사용하고 싶을때!!! 오프라인에서 **페이코**를 사용할때는... Bc카드를 **페이코**를 통해 편의점에서 사용법을 자세하게 알려주세요!! 참고로 기정은...

A **NHN**엔터테인먼트 PAYCO dhk CU,,양사는 제휴를 기념해 1월 1일부터 한달간 도시락 및 일부 스낵을 **페이코**로 결제할 경우 500원을 할인해주는 쿠폰을 제공하며 최대 2...

파워지식iN 답변 pleasure | 쇼핑 > 쇼핑물, 시장 | 1



Q [위비카드요](#) 2017.01.02.

포인트형으로 만들고 싶습니다.

A **NHN** Entertainment 서비스 - 우리카드-박스 제휴이용권 50% 청구할인 - 한국인 출전 **결제** 시 20% 청구할인 - 티켓링크 10% 청구할인 - 코미코 20% 청구할인 - **페이코** 포인트 1...

경제 > 금융 > 신용카드 | 5

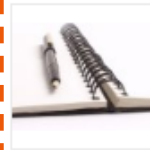


Q [사기당한거 같아요 도와주세요](#) 2016.07.11.

회사는 **NHN**엔터테인먼트, 주식회사 유비핀, playn, 갤럭시마커, 주식회사 마이엠... 보아하니 **페이코**나 이런 **결제**시스템 회사인데 왜 밑도 끝도 없이 이런 큰 돈을...

A 이해가아질않습니다 본인도 모르게 출금이되어다면 회수 **방법**이있습니다 자세한 사항은 쪽지로 연락처 남겨두시면 자세히 상담후 도움드립니다 저희 와이프도 유비핀...

형사사건 > 재산범죄 > 사기 | 2



Q [nhn한국사이버결제 보유중입니다. 어떻게야 할지 속 시원히...](#) 2016.12.06.

**nhn**한국사이버**결제** 보유중입니다. 어떻게야 할지 속 시원히 답변 부탁드립니다. 어찌 하오리까 ??

A 확보중에있구요 **nhn**한국사이버**결제**는 **페이코** 출시 효과에 대한 과도한 시장... 가장 합리적인 주식 투자는 이렇게 올바른 가치 투자 **방법**으로 꾸준히 인내하며 투자를...

경제 > 금융 > 주식, 증권 | 2

# Process

- 문서의 의도를 나타내는 **핵심 구절(Core passage)**을 찾고 이를 통해 문서 간의 유사도를 계산

# Process

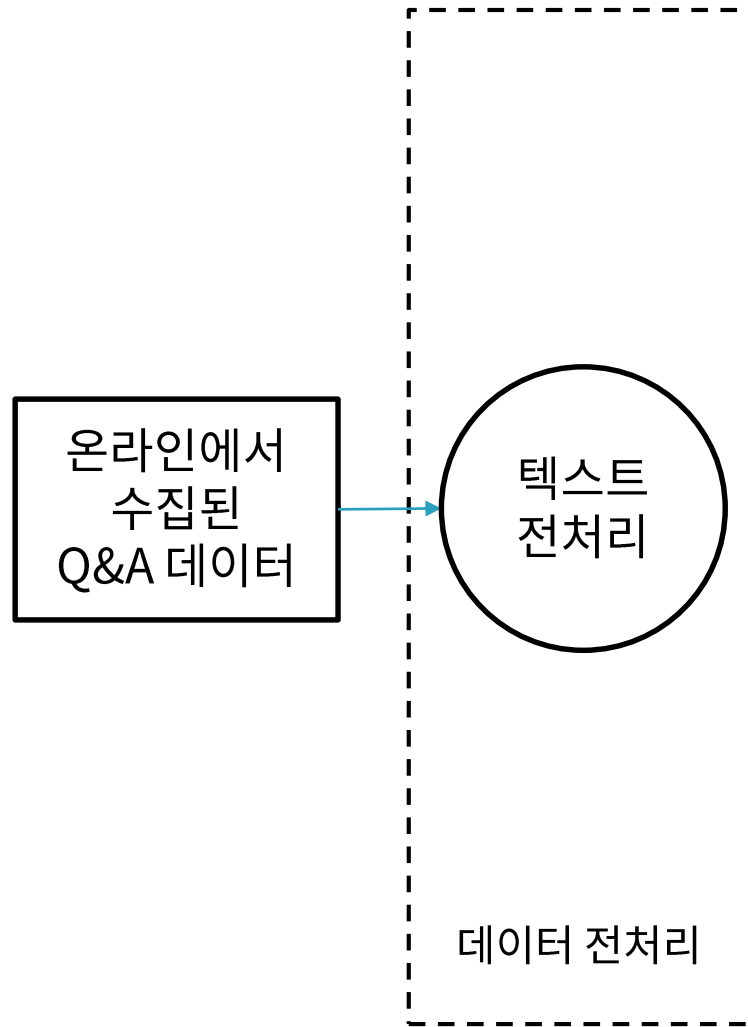
- 문서의 의도를 나타내는 **핵심 구절(Core passage)**을 찾고 이를 통해 문서 간의 유사도를 계산
- 이를 위해 다양한 주제를 가지고 있는 **온라인 Q&A 데이터**를 학습 데이터로 선정

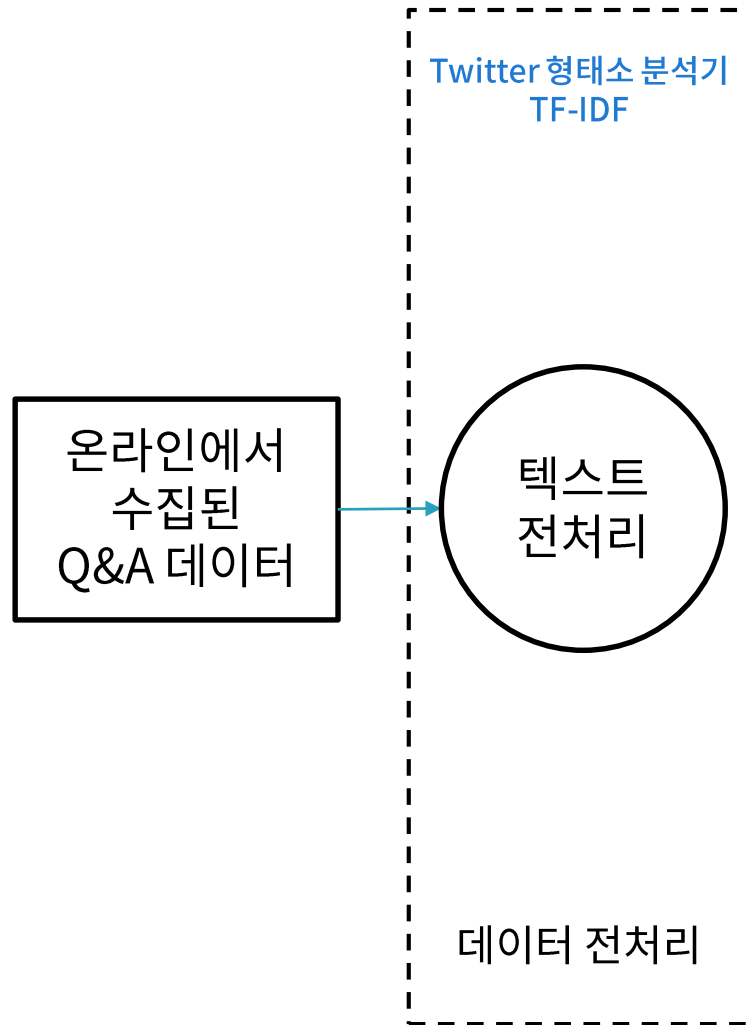
# Process

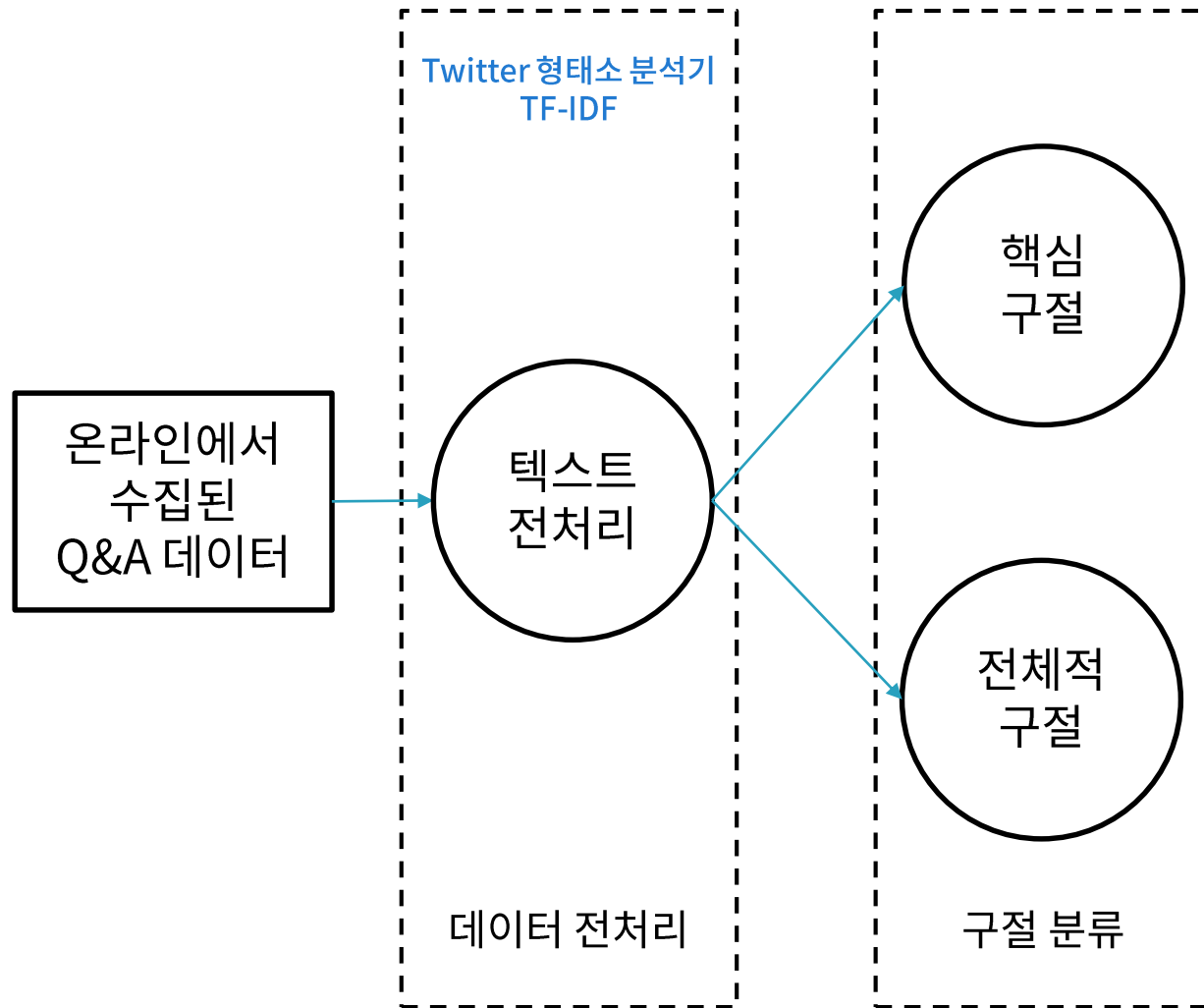
- 문서의 의도를 나타내는 **핵심 구절(Core passage)**을 찾고 이를 통해 문서 간의 유사도를 계산
- 이를 위해 다양한 주제를 가지고 있는 **온라인 Q&A 데이터**를 학습 데이터로 선정
- 이후, 유사한 문서가 얼마나 의도를 반영하는지 **설문조사**를 통해 평가

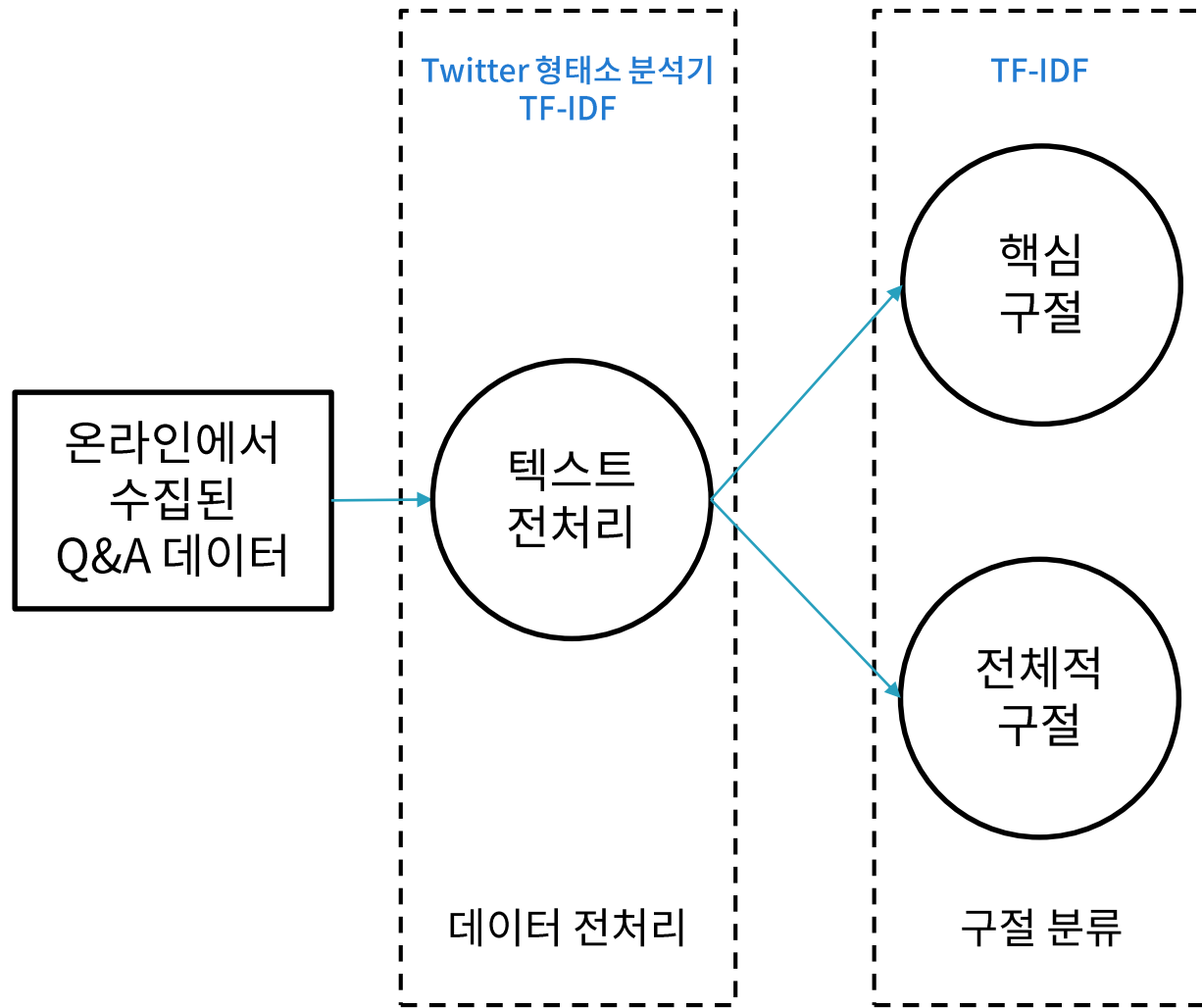
온라인에서  
수집된  
Q&A 데이터

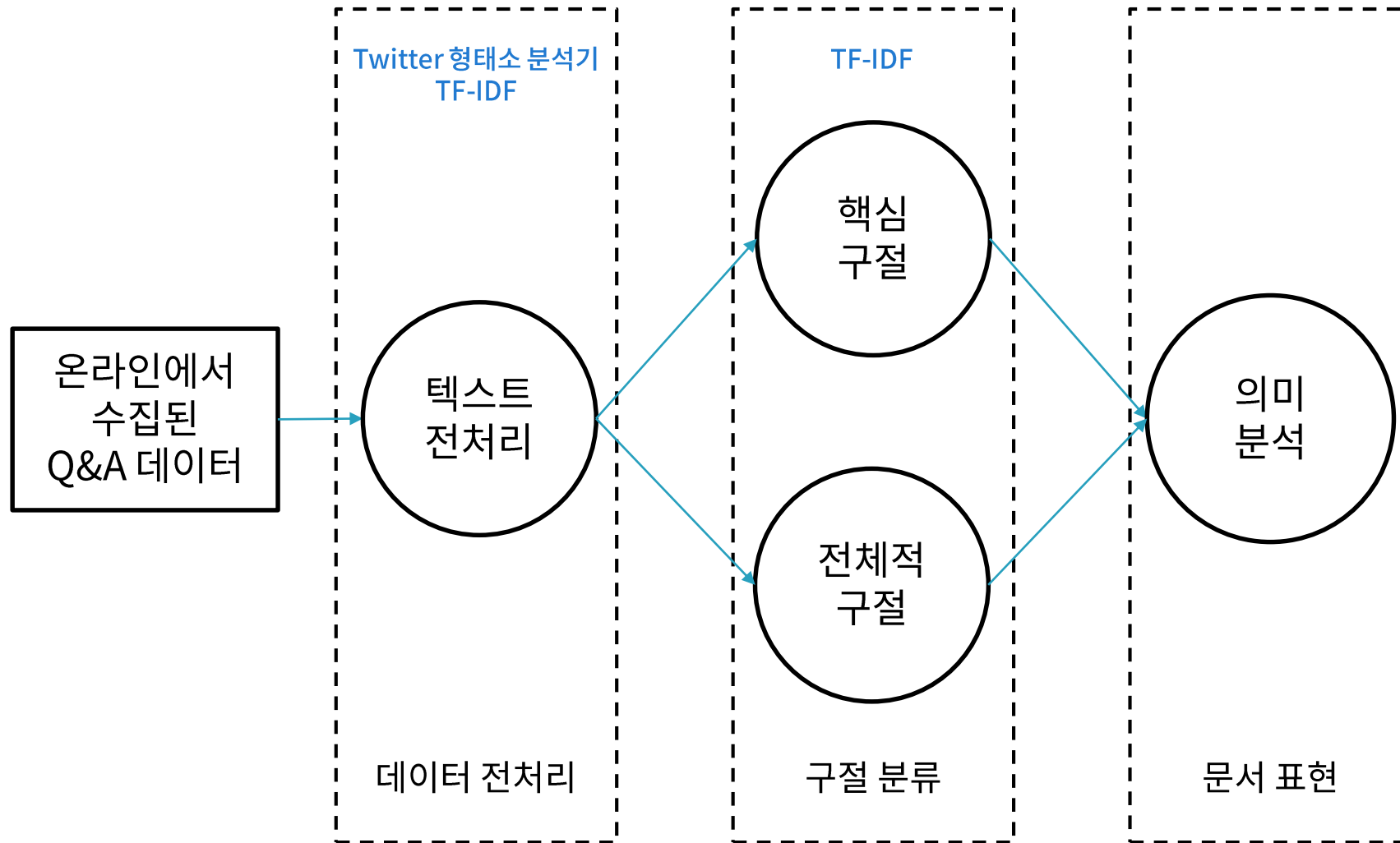


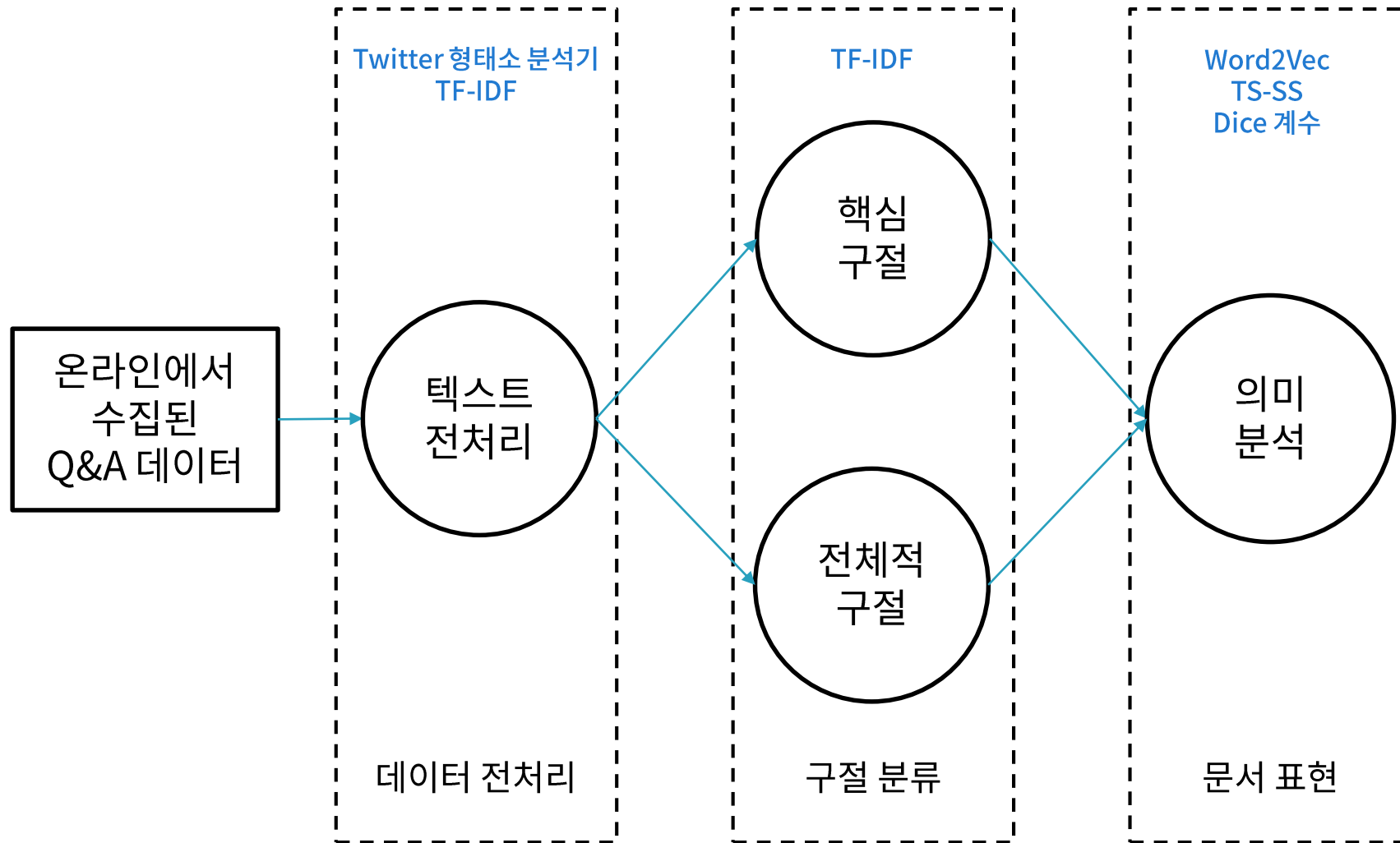


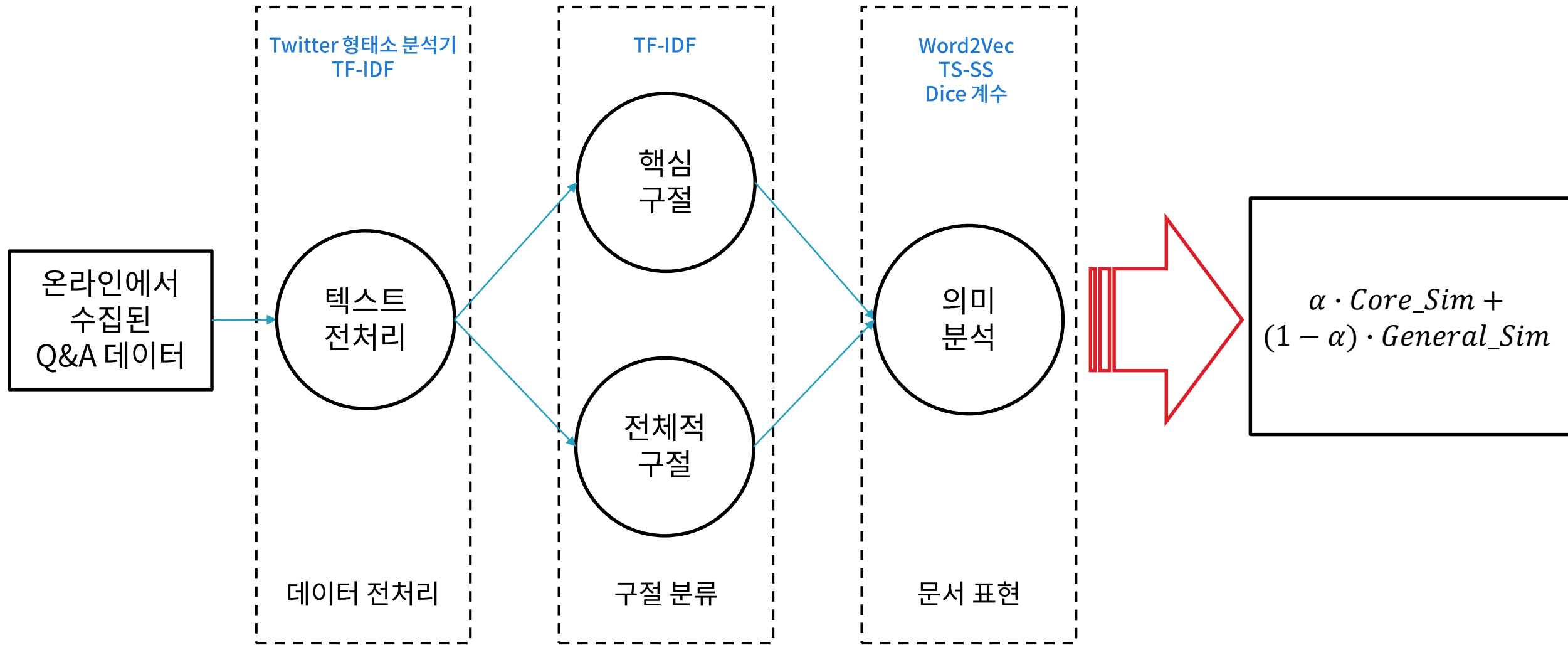












핵심 Idea만  
말해주세요



문장

문장



사람은 무의식적으로 핵심술어를 사용하고, 이 술어 주위로 형성된 어휘들이 의도를 생성한다.

북한이 핵으로 협박하던데, 우리나라랑 전쟁  
날까요? 전쟁나면 어디로 피신해야 하죠?



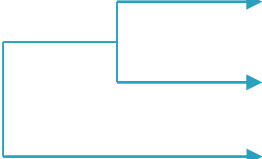

북한이 핵으로 협박하던데, 우리나라랑 전쟁 날까요? 전쟁나면 어디로 피신해야 하죠?	전처리 결과	북한, 핵, 협박, 하다, 우리나라, 전쟁, நா다, 전쟁, நா다, 어디, 피신, 하다, 하다



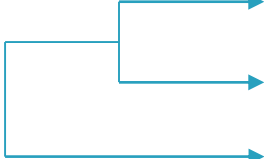
북한이 핵으로 협박하던데, 우리나라랑 전쟁 날까요? 전쟁나면 어디로 피신해야 하죠?	전처리 결과	북한, 핵, 협박, <b>하</b> 다, 우리나라, 전쟁, <b>나</b> 다, 전쟁, <b>나</b> 다, 어디, 피신, <b>하</b> 다, 하다
	핵심 술어	나다, 나다



북한이 핵으로 협박하던데, 우리나라랑 전쟁 날까요? 전쟁나면 어디로 피신해야 하죠?	전처리 결과	북한, 핵, 협박, <b>하</b> 다, 우리나라, 전쟁, <b>나</b> 다, 전쟁, <b>나</b> 다, 어디, 피신, <b>하</b> 다, 하다
	핵심 술어	나다, 나다
	핵심 명사	(우리나라,전쟁,전쟁), (전쟁,어디,피신)

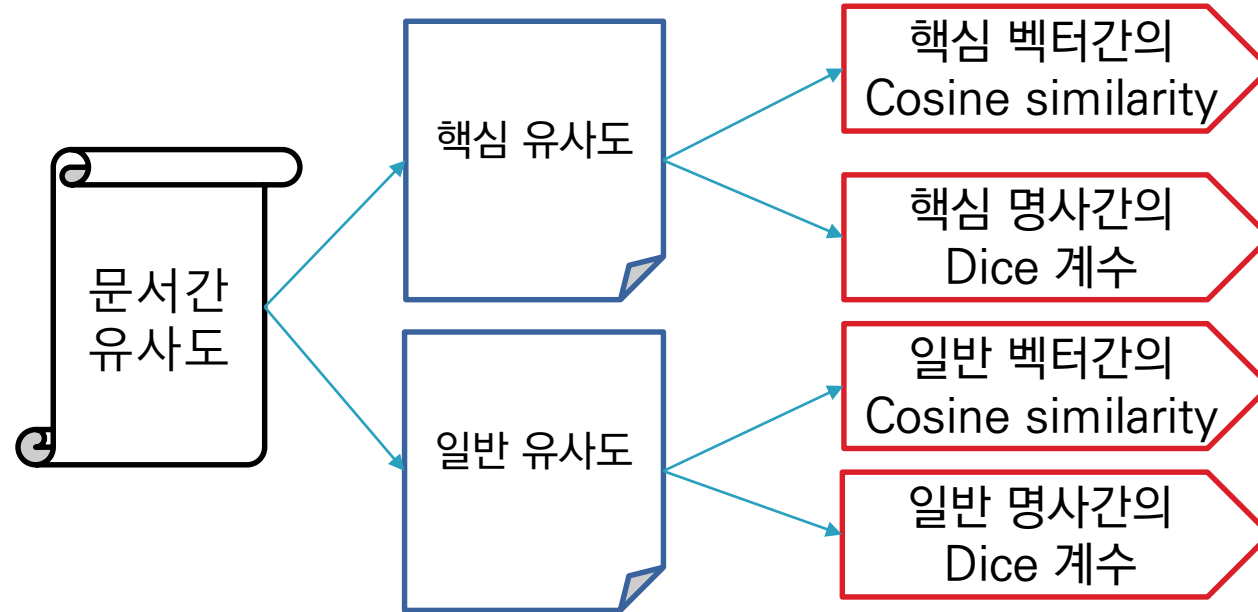
북한이 핵으로 협박하던데, 우리나라랑 전쟁 날까요? 전쟁나면 어디로 피신해야 하죠?  	전처리 결과	북한, 핵, 협박, <b>하</b> 다, 우리나라, 전쟁, <b>나</b> 다, 전 쟁, <b>나</b> 다, 어디, 피신, <b>하</b> 다, 하다
	핵심 술어	나다, 나다
	핵심 명사	(우리나라,전쟁,전쟁), (전쟁,어디,피신)
	핵심 구절	우리나라, 전쟁, 나다, 전쟁, 전쟁, 나다, 어디, 피신



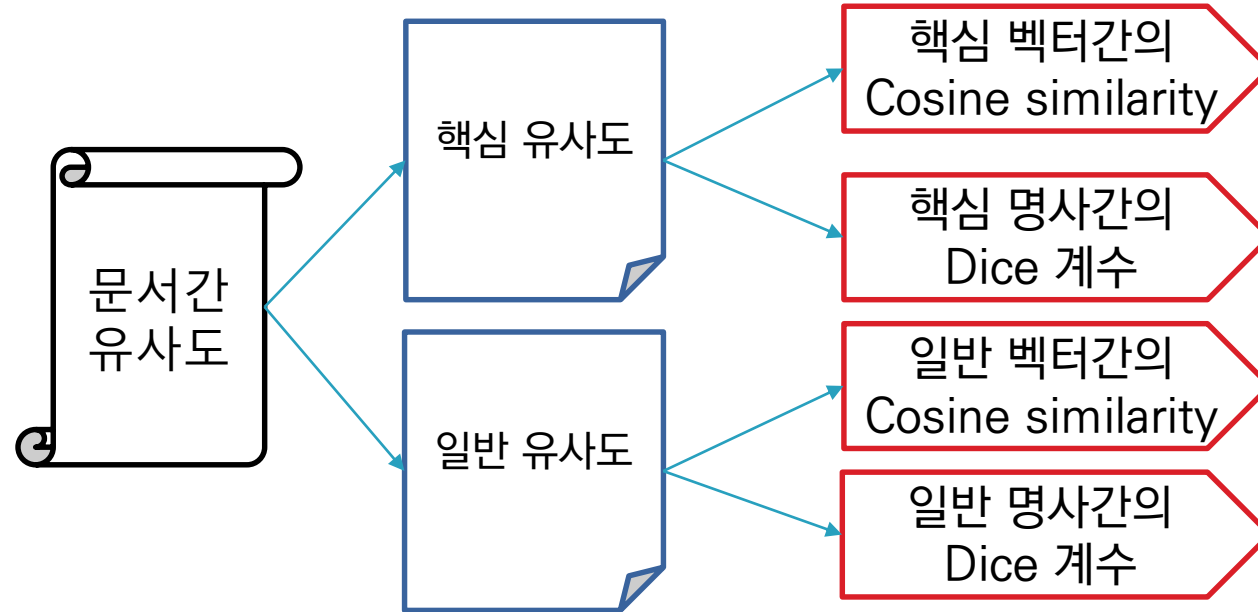
<p>북한이 핵으로 협박하던데, 우리나라랑 전쟁 날까요? 전쟁나면 어디로 피신해야 하죠?</p> <div></div>	전처리 결과	북한, 핵, 협박, <b>하</b> 다, 우리나라, 전쟁, <b>나</b> 다, 전쟁, <b>나</b> 다, 어디, 피신, <b>하</b> 다, 하다
	핵심 술어	나다, 나다
	핵심 명사	(우리나라,전쟁,전쟁), (전쟁,어디,피신)
	핵심 구절	우리나라, 전쟁, 나다, 전쟁, 전쟁, 나다, 어디, 피신
	전체적 구절	북한, 핵, 협박, 우리나라, 전쟁, 어디, 피신



<p>북한이 핵으로 협박하던데, 우리나라랑 전쟁 날까요? 전쟁나면 어디로 피신해야 하죠?</p>	전처리 결과	북한, 핵, 협박, <b>하</b> 다, 우리나라, 전쟁, <b>나</b> 다, 전쟁, <b>나</b> 다, 어디, 피신, <b>하</b> 다, 하다
	핵심 술어	나다, 나다
	핵심 명사	(우리나라,전쟁,전쟁), (전쟁,어디,피신)
	핵심 구절	우리나라, 전쟁, 나다, 전쟁, 전쟁, 나다, 어디, 피신
	전체적 구절	북한, 핵, 협박, 우리나라, 전쟁, 어디, 피신



<p>북한이 핵으로 협박하던데, 우리나라랑 전쟁 날까요? 전쟁나면 어디로 피신해야 하죠?</p>	전처리 결과	북한, 핵, 협박, <b>하</b> 다, 우리나라, 전쟁, <b>나</b> 다, 전쟁, <b>나</b> 다, 어디, 피신, <b>하</b> 다, 하다
	핵심 술어	나다, 나다
	핵심 명사	(우리나라,전쟁,전쟁), (전쟁,어디,피신)
	핵심 구절	우리나라, 전쟁, 나다, 전쟁, 전쟁, 나다, 어디, 피신
	전체적 구절	북한, 핵, 협박, 우리나라, 전쟁, 어디, 피신

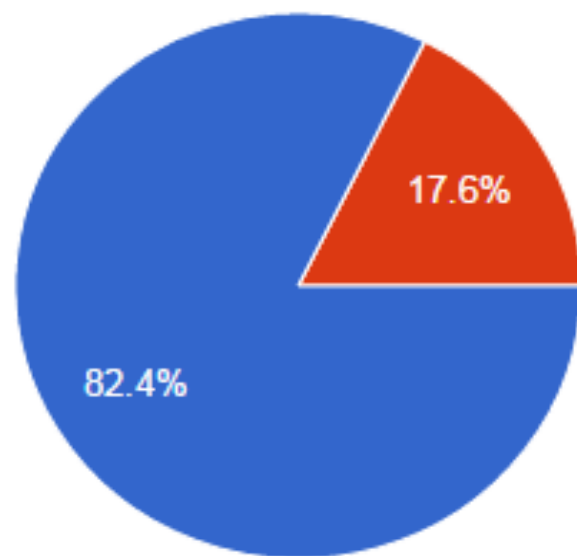


$$\alpha \cdot Core\_Sim + (1 - \alpha) \cdot General\_Sim$$

결과는 ?

제가학교에서ttp자료를조사해서하는건데 제가세월호에대해발표를하는데 세월호 사망자 생존자 실종자 수를 빨리보고해야되는데 빨리답변부탁드립니다.

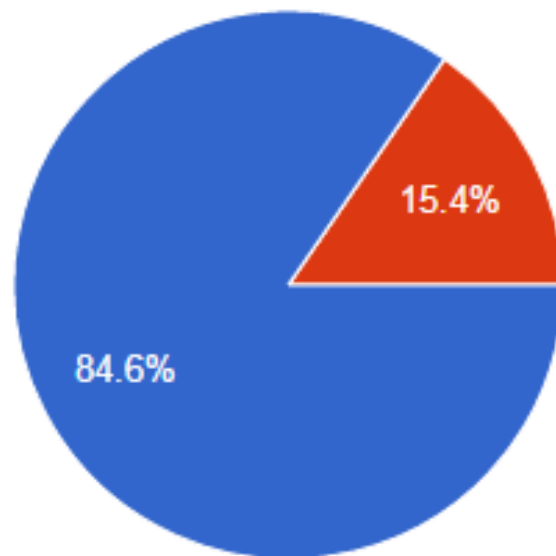
(응답 102개)



- 1.세월호 실종자 수 좀 알려주시고요  
2.사망자수 도 알려주시고 3.생존자수 도 좀 알려주세요!! 내공 100입니다
- 학교에서 제가 자료조사 숙제 맡았거든요. 생존자나 사망자 등 세월호에 대해 다 알려주세요

막 아까 뉴스 봤는데 북한이 핵탄두 한다던데... 핵을 쏜다는건가요? 핵을쏘면 전쟁하는건데 전쟁나요? 전쟁날확률ㅠㅠ내공100

(응답 104개)



- 이질문에 답해주세요! 1.북한전쟁 일어나나요? 2.핵쏘나요? 3.전쟁일어날 가능성은몇프로인가요?
- 방금 뉴스에서 북한과 전쟁할수도 있다는데 만약 전쟁을 한다면 어떻게 해야 할까요? 그리고 좀 이상한 말 같겠지만 지하로 가면 핵을 피할수 있을까요?ㅋ..... 답변 부탁드립니다..

이제 더 없나요?



# Another Project / Python, Tensorflow

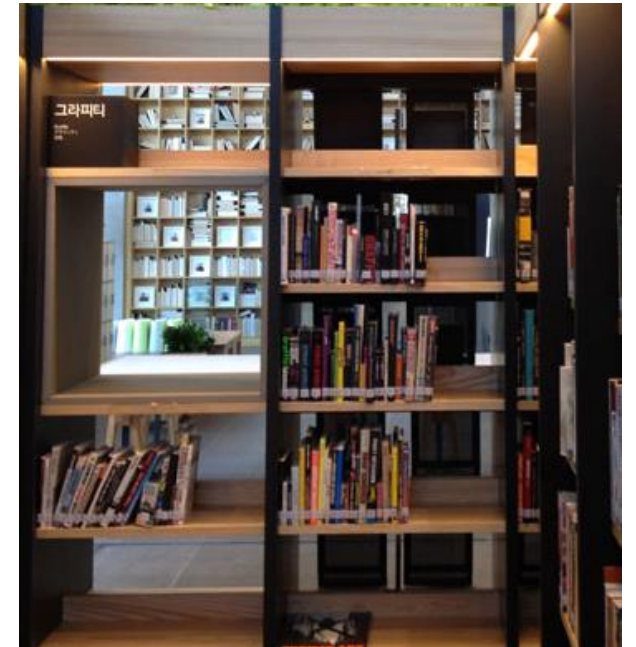
## Another Project / Python, Tensorflow

인공신경망이 핫하던데  
Python이랑 Tensorflow  
공부해보고 싶다...

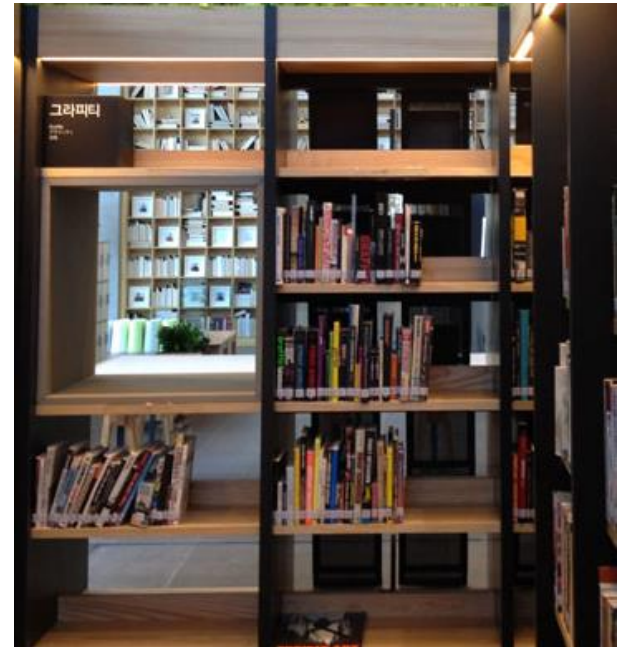
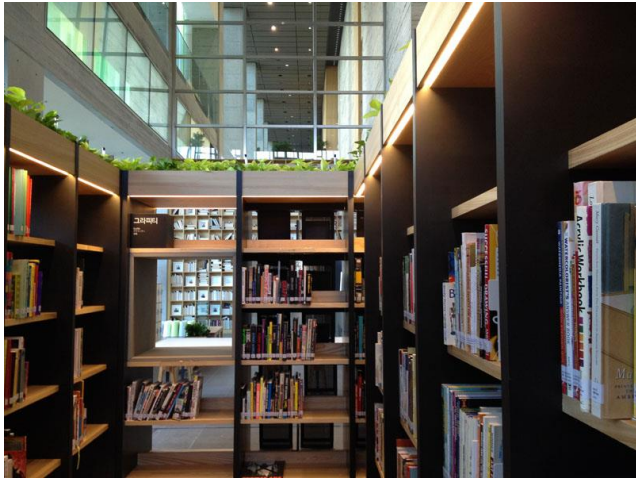




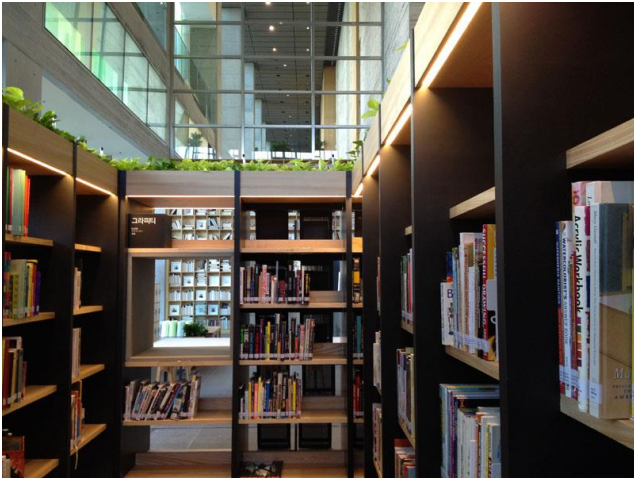






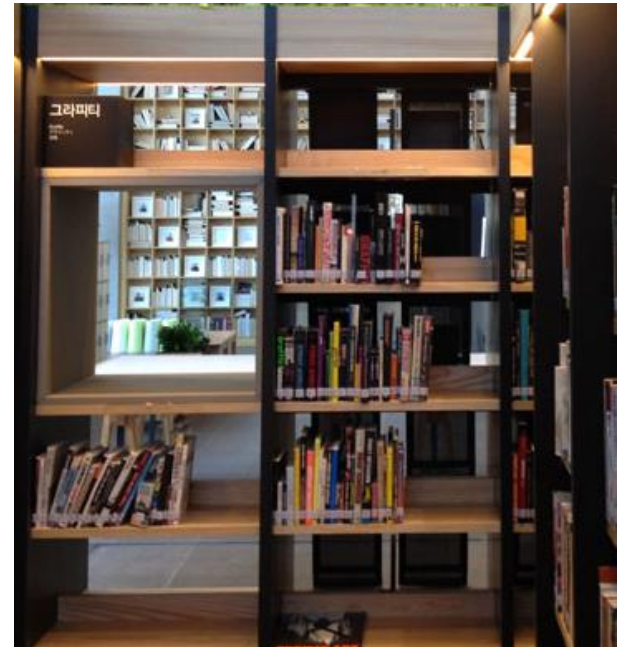
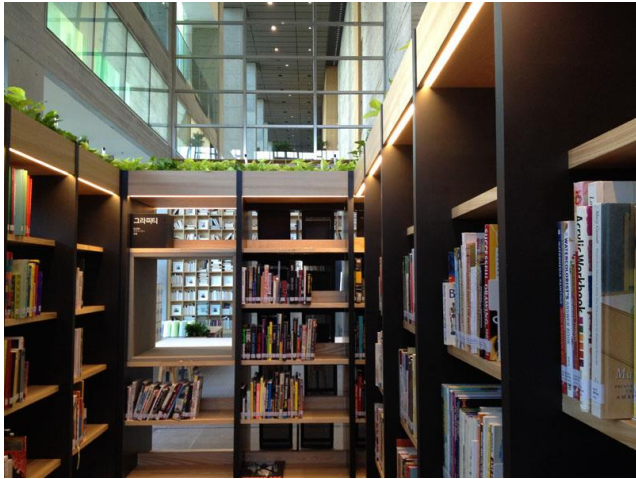


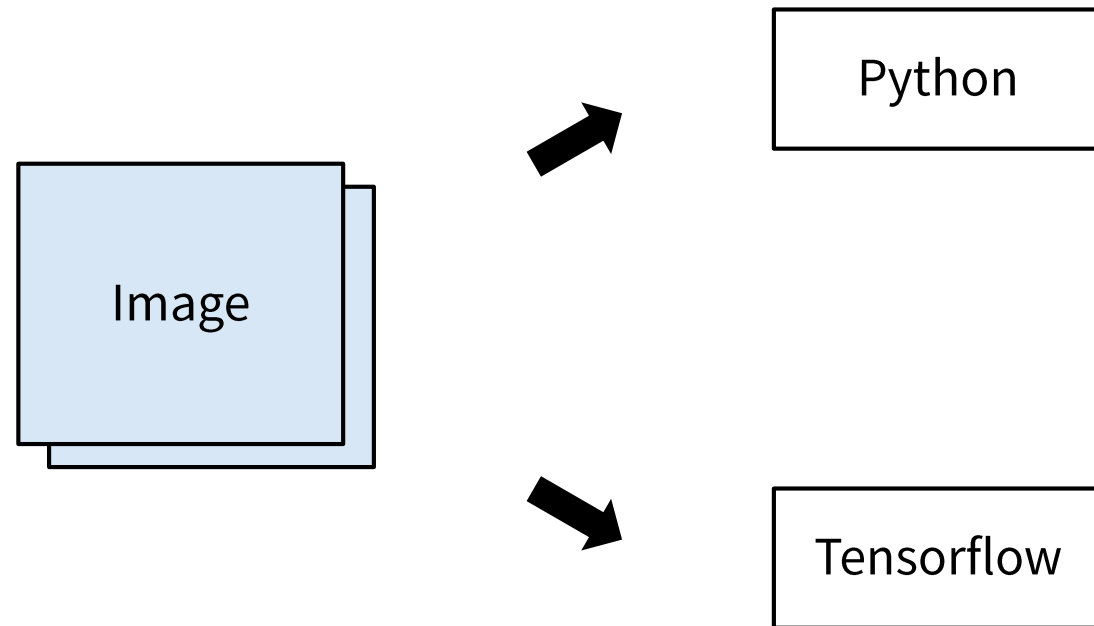
# 동일 이미지

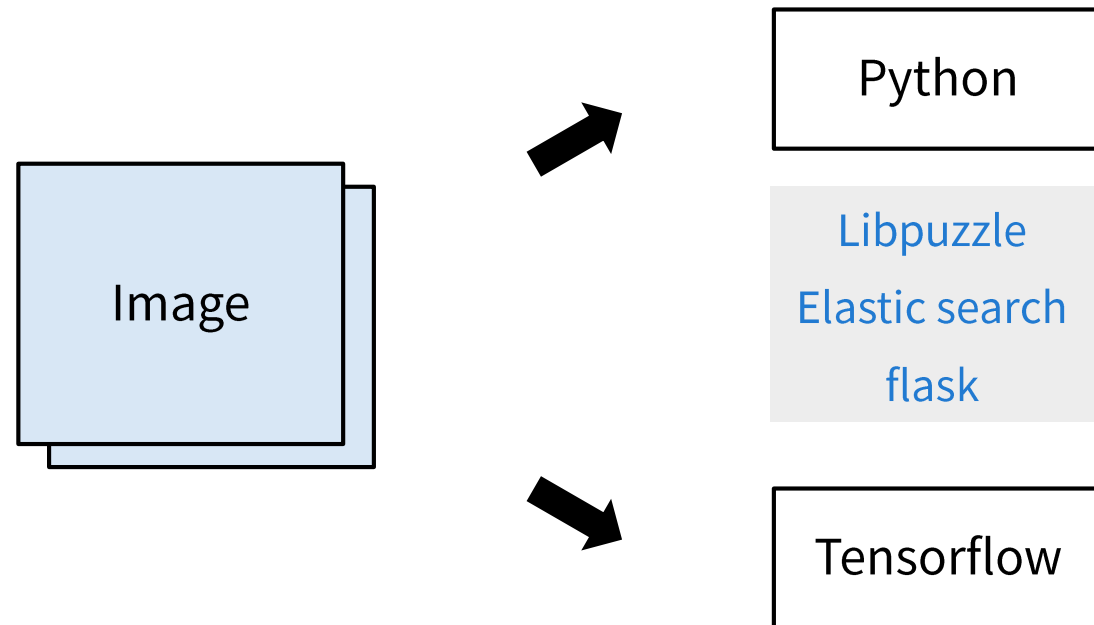




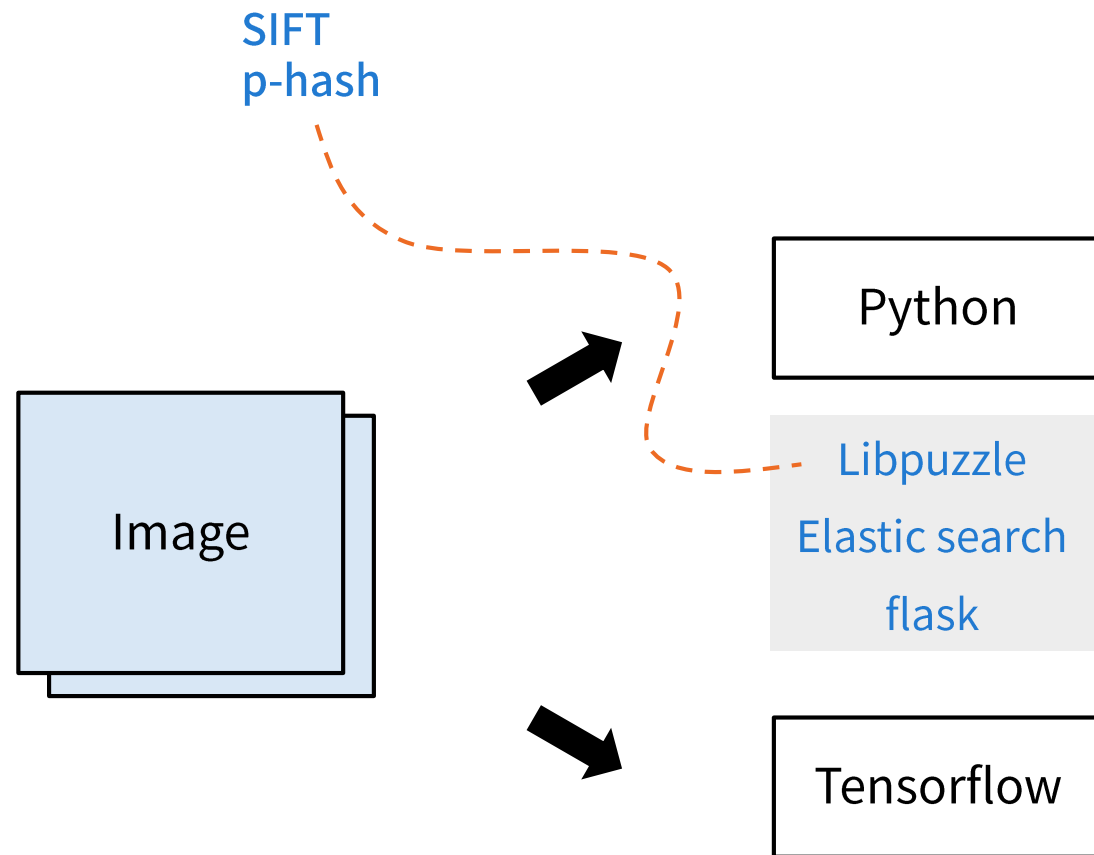
# 동일 카테고리

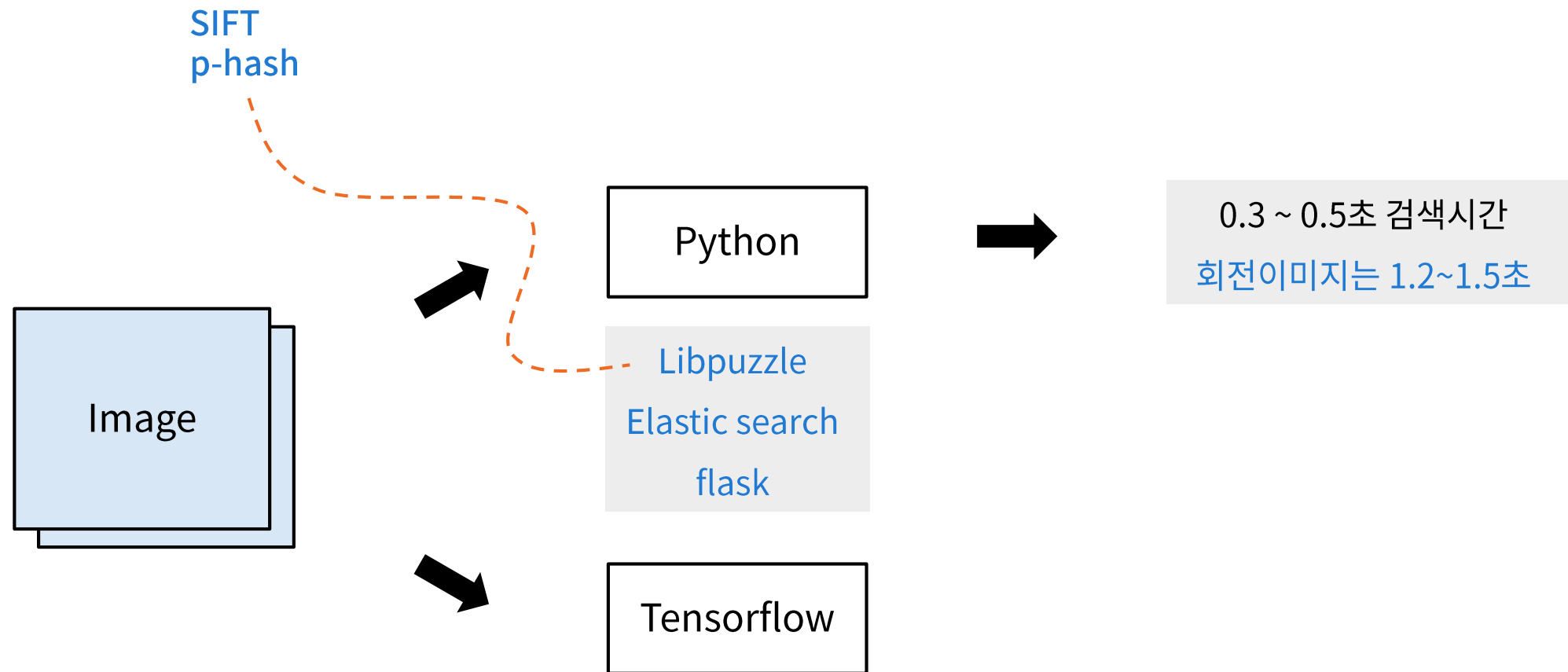


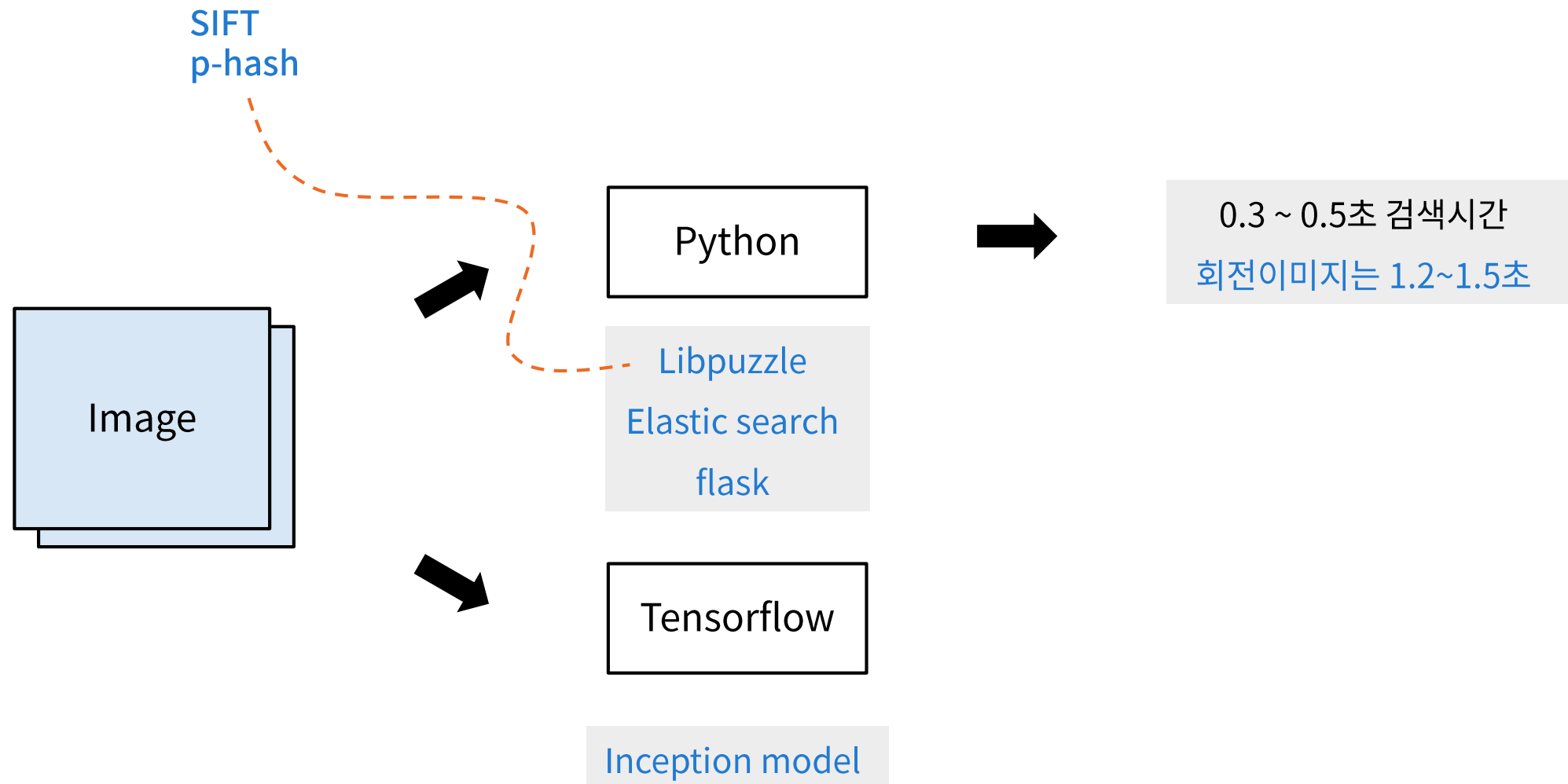


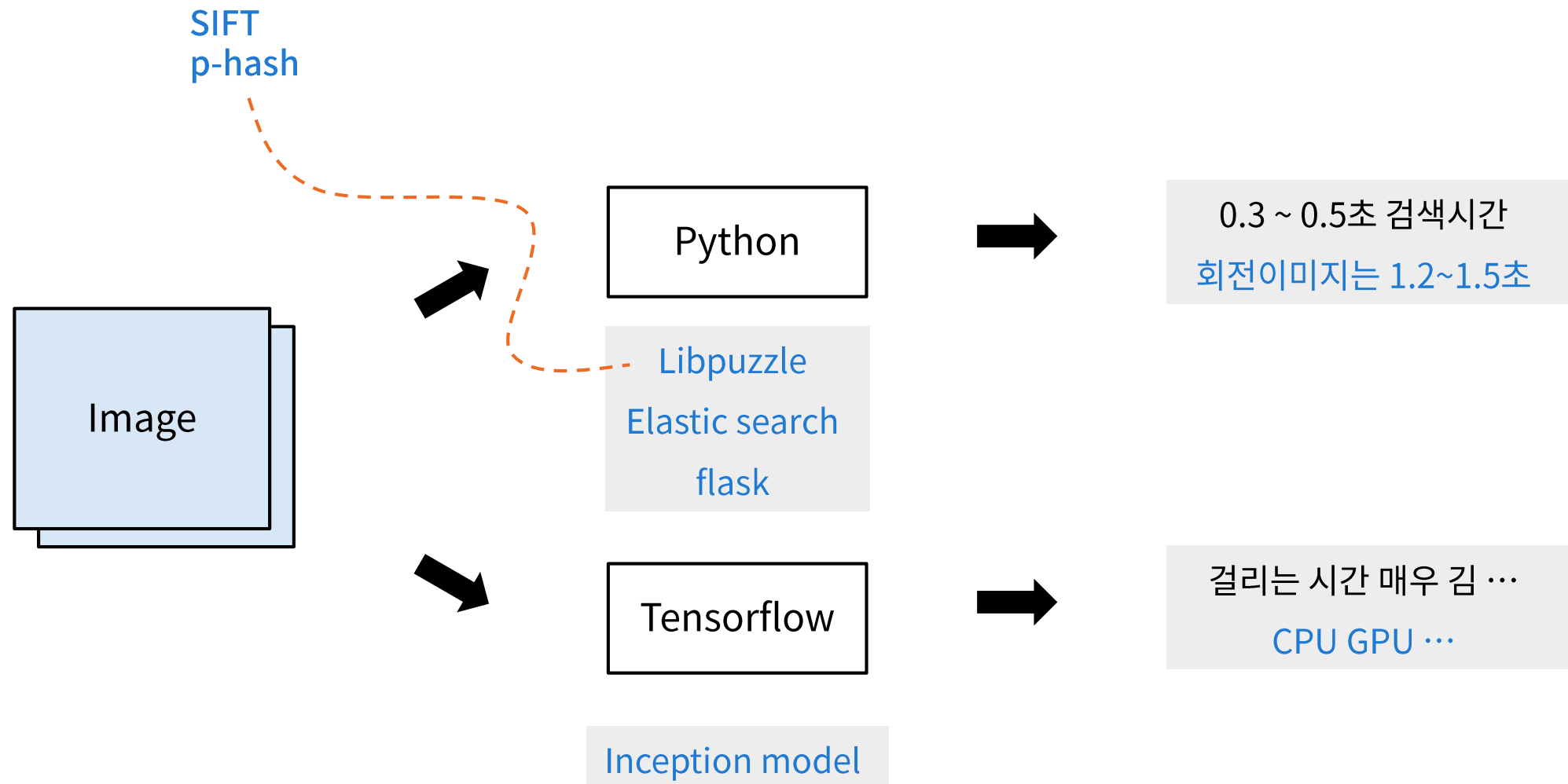












# 하지 못했던 말

법률바다  
2017년 서비스 예정

이외에도  
법률 로보 어드바이저

# 하지 못했던 말

법률바다  
2017년 서비스 예정

AIP Proceeding conference

이외에도

법률 로보 어드바이저

일베와 오유 데이터를 이용한 성향 분류

# 하지 못했던 말

법률바다  
2017년 서비스 예정

AIP Proceeding conference

이외에도

법률 로보 어드바이저

일베와 오유 데이터를 이용한 성향 분류

WordEmbedding 기법을 이용한 Lotto 예측...

혼자 연구중인 프로젝트

이제 더 없나요?





하.. 한마디만...

믿음을 주시면,  
그 믿음에 보답하겠습니다.

Thank you