

Part 3. 프롬프트 엔지니어링 개요

검색엔진의 진화 Vector Search





유연 근무제가 있는
이커머스 플랫폼 개발 포지션 찾아줘

이커머스 플랫폼 개발

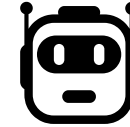
플래티어 | 응답률 평균 이상 | 서울 · 한국

#인원급성장 #51-300명 #설립10년이상 #유연근무 #식비 #음료 #커피 #와인
#IT, 컨텐츠

- 일과 삶의 균형도 중요하죠
- 빠른데이 (매월 셋째 주 수요일 5시 퇴근) 문화가 있어요
- **유연 근무제(8시반 ~ 10시반 자율 선택)를 시행하고 있어요**
- 설날, 추석, 창립기념일은 따뜻하게 보내세요(상여급 지급)
- 자유로운 연차 사용을 보장해요.



제공한 포지션 정보를 기반으로,
사용자가 요청한 내용에 대해 요약과
추천사를 작성해줘



유연 근무제가 있는 이커머스 플랫폼 회사로
플래티어의 플랫폼 개발자 포지션을 찾았습니다.

플래티어는 8시반~10시반 자율 선택으로
유연 근무제를 시행하고 있으며, 자유로운 연차
사용 보장으로 근무를 유연하게 할 수 있는
환경을 갖춘 회사여서 적극 추천합니다.

Personalized Answer Generation



유연 근무제가 있는
이커머스 플랫폼 개발 포지션 찾아줘

이커머스 플랫폼 개발

플래티어 [응답률 평균 이상](#) | 서울 · 한국

#인원급성장 #51~300명 #설립10년이상 #유연근무 #식비 #음료 #커피 #와인
#IT, 콘텐츠

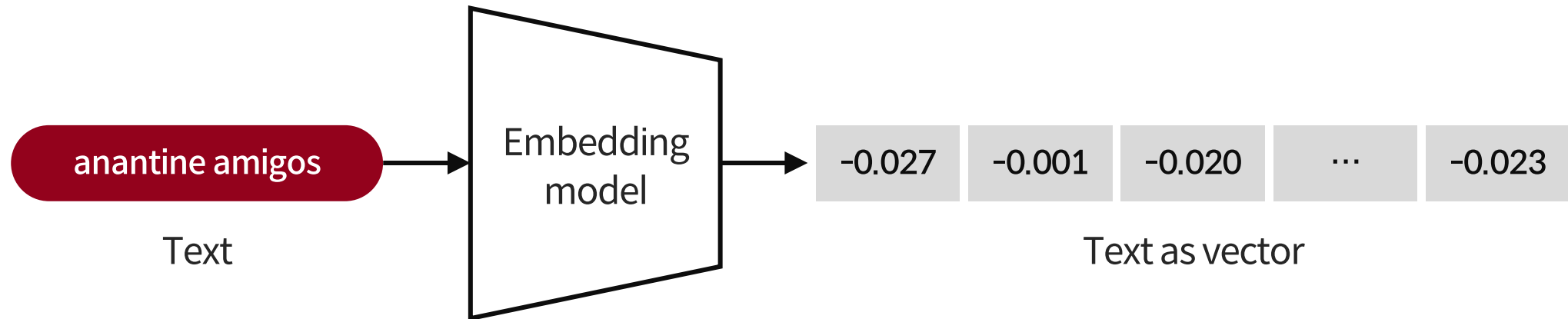
- 일과 삶의 균형도 중요하죠
- 빠른데이 (매월 셋째 주 수요일 5시 퇴근) 문화가 있어요
- **유연 근무제(8시반 ~ 10시반 자율 선택)를 시행하고 있어요**
- 설날, 추석, 창립기념일은 따뜻하게 보내세요(상여급 지급)
- 자유로운 연차 사용을 보장해요.

← 어떻게 찾아오지?

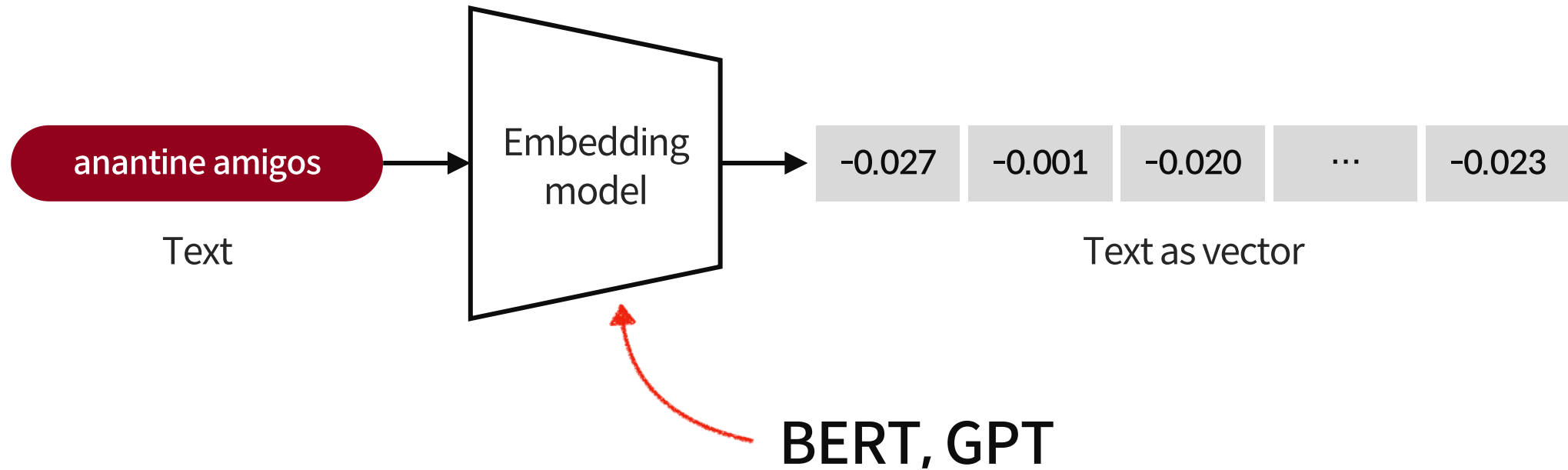
LLM의 숨은 영웅 Embedding



Embedding



Embedding



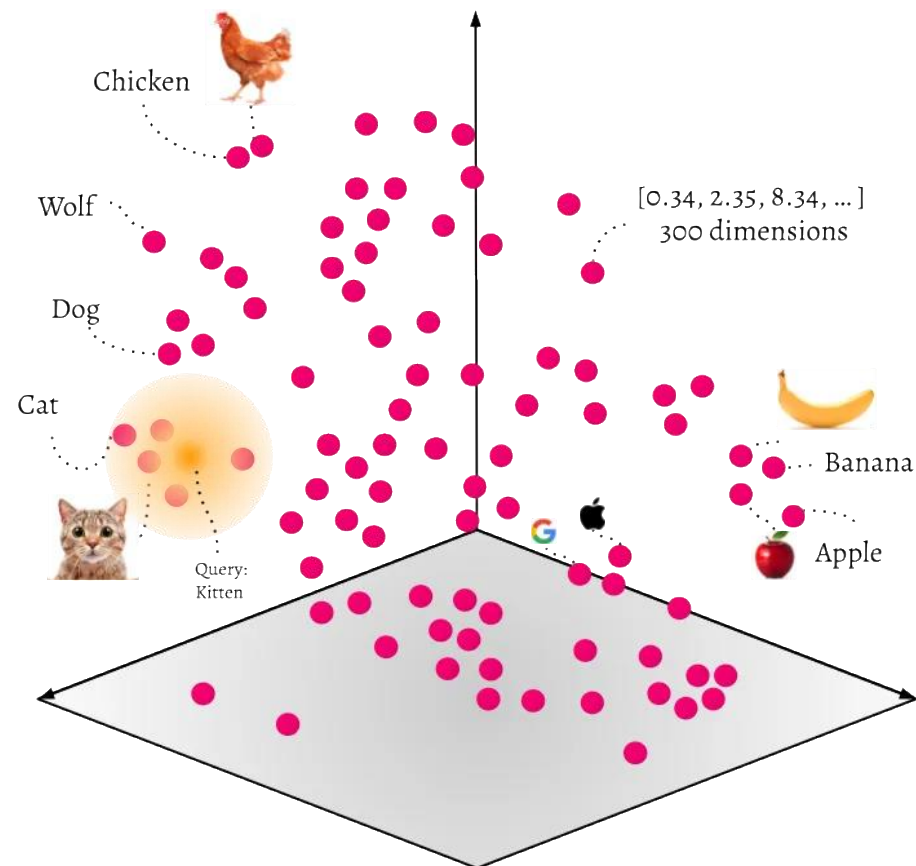
Embedding Vector

Word	Embedding Vector			
Apple	-0.82	-0.32	...	-0.23
House	0.419	1.28	...	-0.06
Car
Orange	-0.74	-1.02	...	1.35

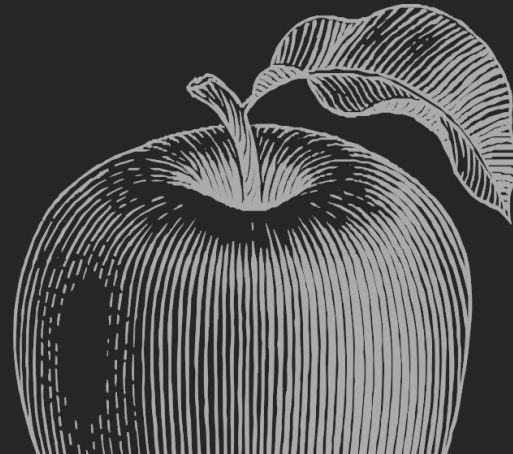
Embedding Space

	Apple			
	Banna			
			House	
				Car

Embedding Space



Vector Search & Semantic Search



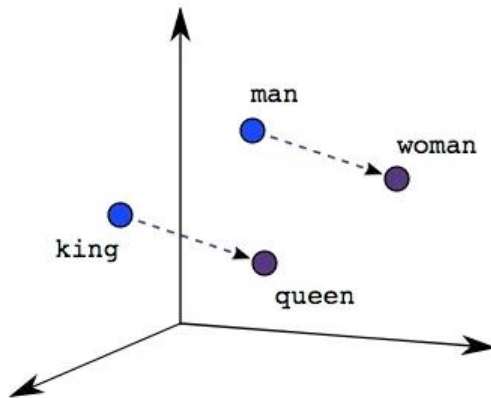
Vector Search

	Apple			
	Banna			
			House	
				Car

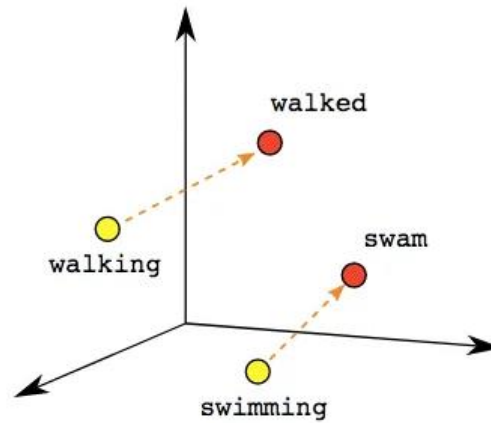
Vector Search

	Apple	<div>Distance or Similarity</div>		
	Banna			
			House	
				Car

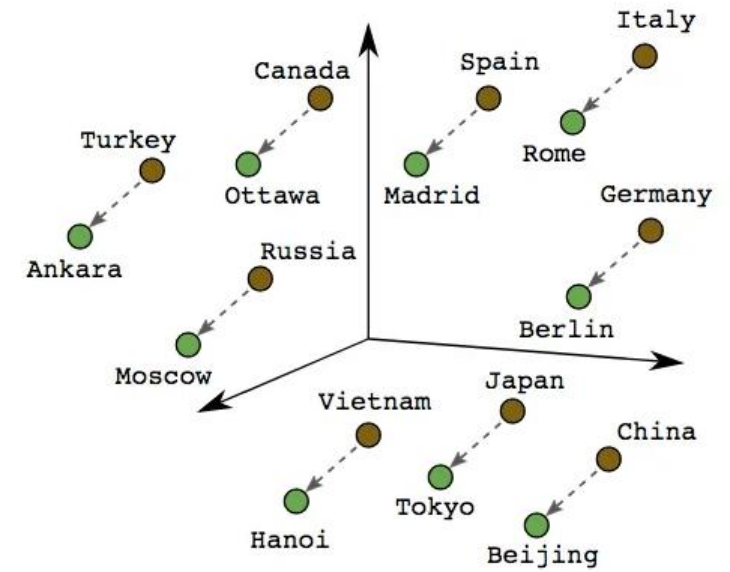
Vector Search



Male-Female



Verb Tense



Country-Capital

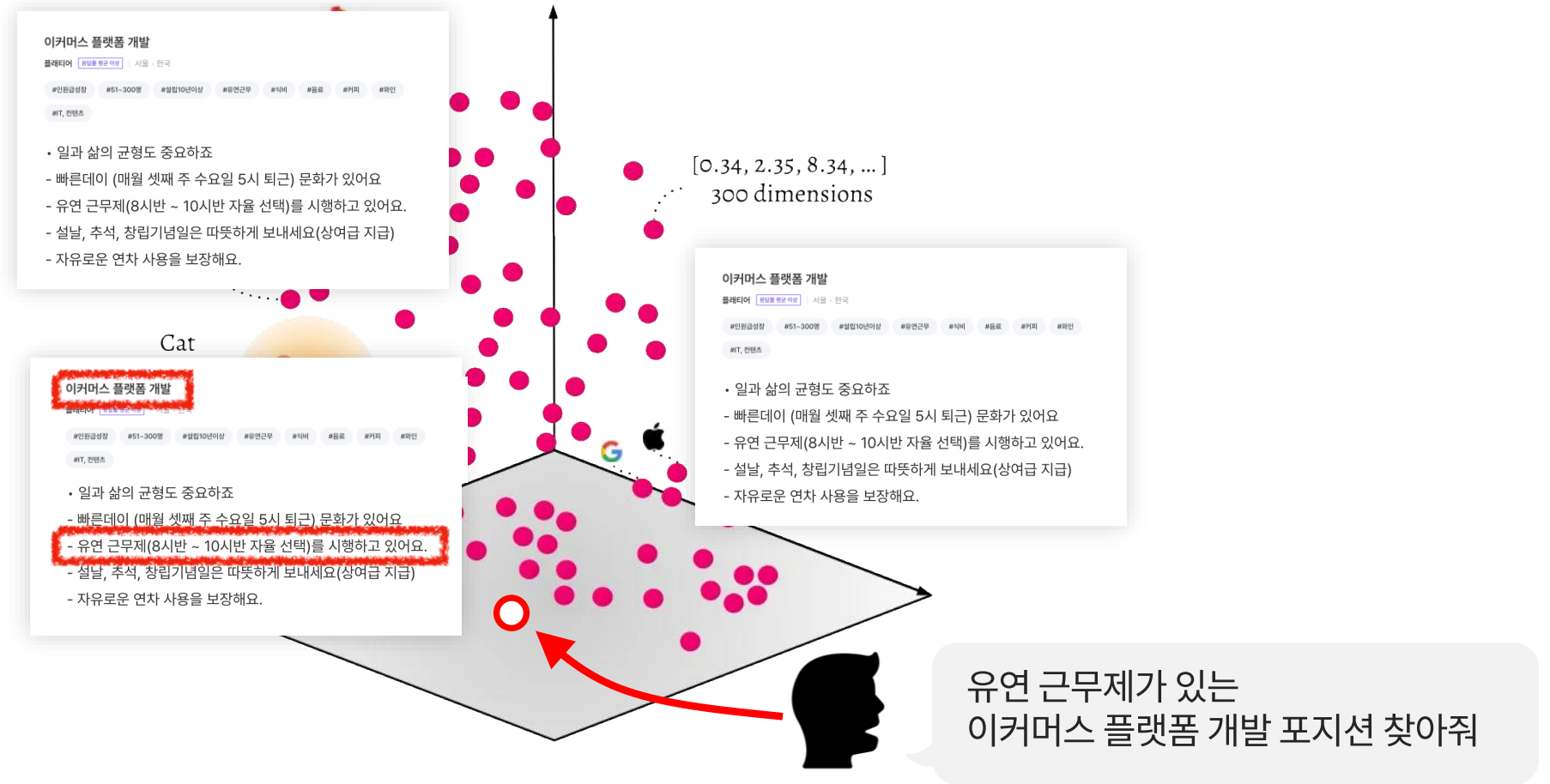
Sentence Embedding

Sentence	Embedding Vector			
Hello, how are you?	-0.82	-0.32	...	-0.23
It's a beautiful day	0.419	1.28	...	-0.06
...
I'm fine, and you?	-0.74	-1.02	...	1.35

Embedding Space

	Hello, How are you?			
I'm fine, and you?				
		It's a beautiful day		
				...

Vector Search → Semantic Search



Why?

No more 풀 텍스트 인덱싱

No more 형태소 분석

No more 유사어/다국어 처리

No more 분산 시스템

...

Performance

Text search Code search Sentence similarity Text classification

Model	Performance
text-embedding-ada-002	81.5
text-similarity-davinci-001	80.3
text-similarity-curie-001	80.1
text-similarity-babbage-001	80.1
text-similarity-ada-001	79.8

Dataset: [SentEval](#) (STS 2012–2016)

COST

Wikipedia 전체 콘텐츠 색인 = 100만원

Wikipedia 전체 콘텐츠 생성 = 543만원

글 6백만개 기준 ≍ 일년간 생산된 한국 전체 뉴스 기사

COST

Wikipedia 전체 콘텐츠 색인 = 100만원

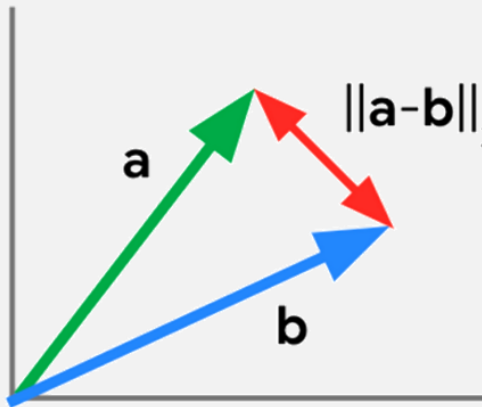
Wikipedia 전체 콘텐츠 생성 = 543만원

전세계 사람 대상의 개인화된 백과사전 개발 비용 = 100만원

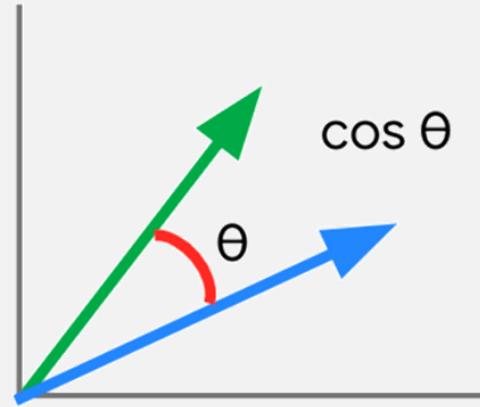
But, 만능은 아님

속도 문제

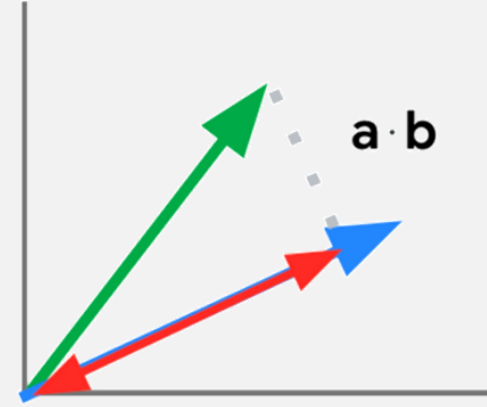
Brute force search takes too long: $O(\text{dims} \times \text{items})$



L2 distance



Cosine similarity

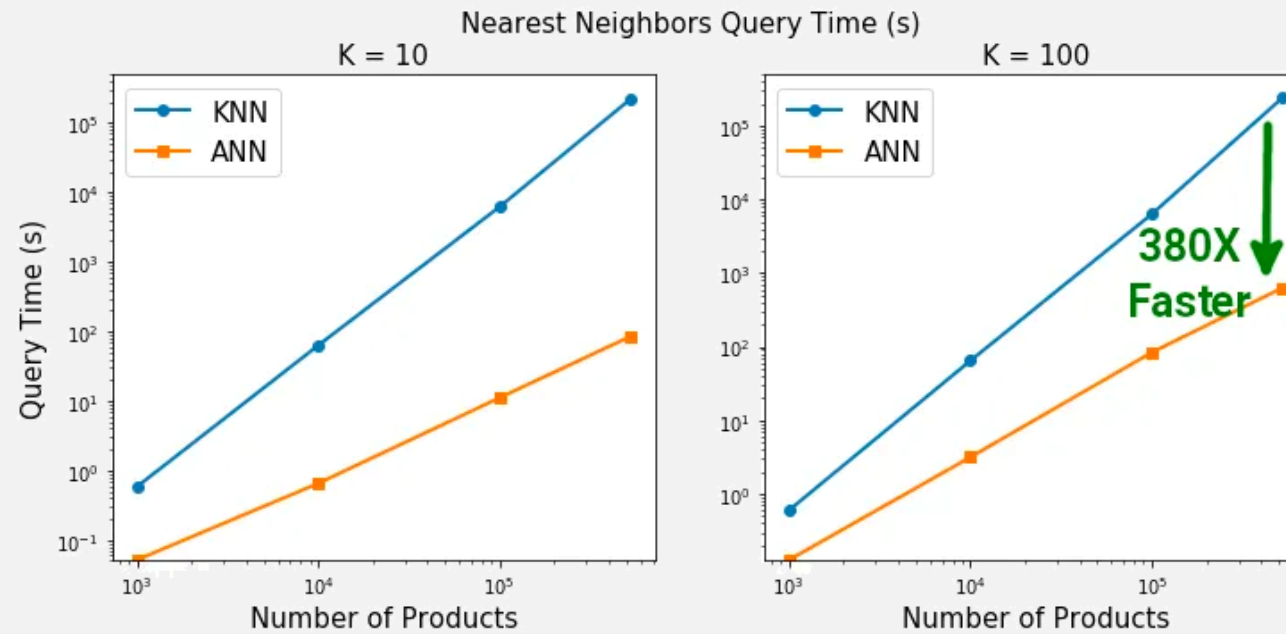


Inner product

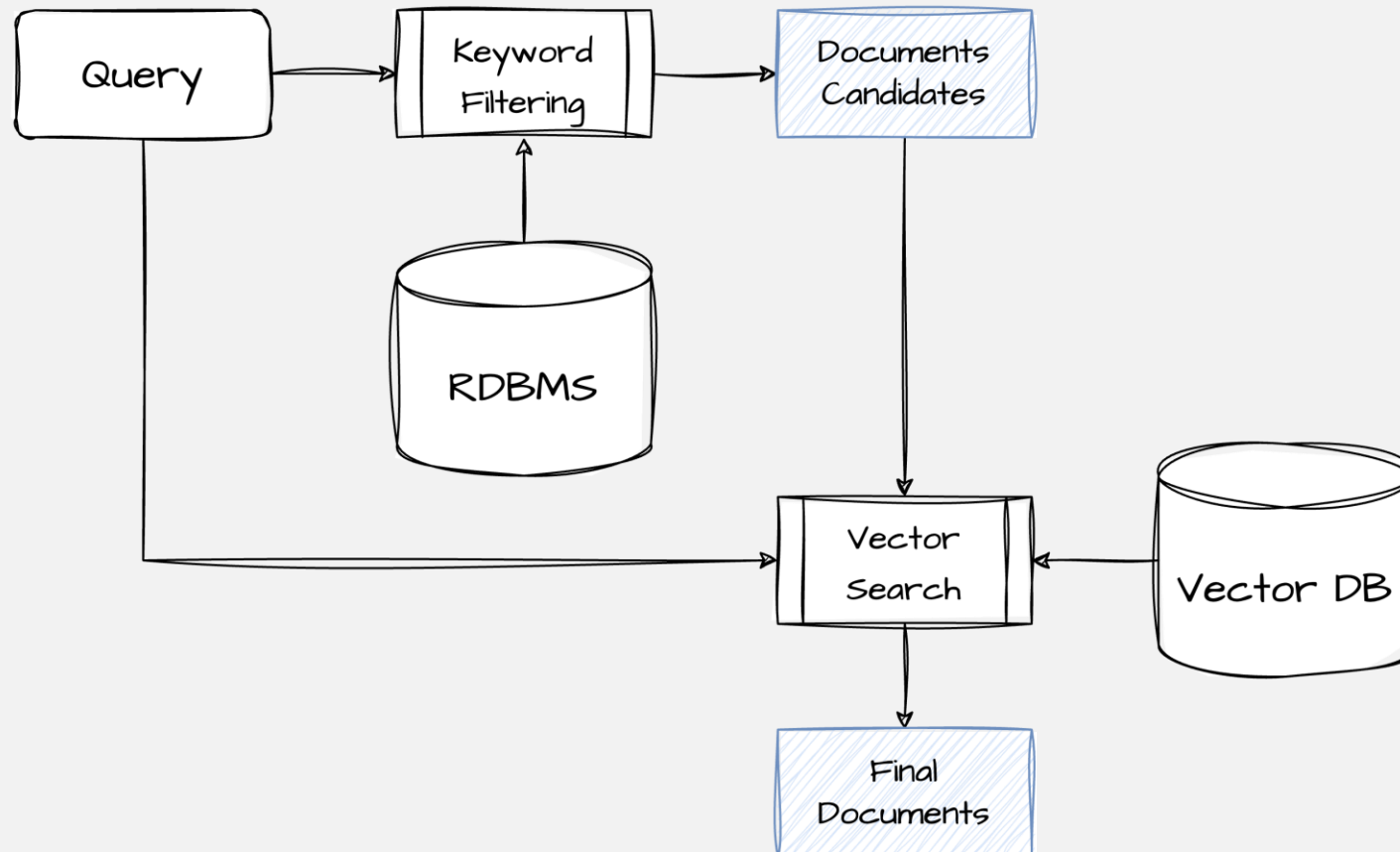
성능 문제

ANN (Approximate Nearest Neighbor) 사용

속도 문제로 모든 벡터를 비교하지 않고, 근사값을 찾는 방법



Hybrid Search



Prompt Engineering

In-Context Learning

- 답변을 위해 필요한 적절한 컨텍스트 제공
- 원하는 결과를 추출을 위한 프롬프트 작성
- 원하는 포맷의 출력을 위한 프롬프트 작성

Reasoning chain #1

Q: What is Brad Peyton's occupation?

A: Brad Peyton is a film director.

Q: What is seismology?

A: Seismology is the study of earthquakes.

Q: Do film directors need to know about earthquakes?

A: No, film directors do not need to know about earthquakes.

Reasoning chain #2

Q: What did Brad Peyton direct?

A: Brad Peyton is a Canadian film director.

...

Q: What is San Andreas about?

A: San Andreas is a film directed by Brad Peyton, about a massive earthquake caused by the San Andreas Fault.

Reasoning chain #3

Q: What is Brad Peyton's occupation?

A: Brad Peyton is a film director, writer, and producer.

Q: What do film directors have to know?

A: Film directors have to know a lot of things.

Q: Is seismology one of the things that Brad Peyton has to know?

A: No.

Multi-Chain Reasoning (MCR)



Answer: Brad Peyton is a film director, writer, and producer. Seismology is the study of earthquakes. San Andreas is a film directed by Brad Peyton, about a massive earthquake caused by the San Andreas Fault.

So the final answer is: Yes.

