## 랭체인(LangChain)으로 AI 웹 서비스 만들기

- 조코딩 -

### 목차

- 1. 강의 개요
- 2. 개념 설명
- 3. 환경 세팅
- 4. LangChain 시작하기
- 5. 인공지능 시인 만들기
- 6. ChatPDF 만들기

# 강의 개요

#### 이번 강의로 만들 것





#### 인공지능 시인

어떤 것에 관한 시를 써드릴까요?

조코딩

시의뢰하기

조코딩은 즐거운 여행이야 키보드와 마우스로 더해지는 마법 코드 한 줄이 세상을 바꿀 수 있어 프로그 래머의 손에 비상한 힘이 깃들어

세미콜론과 괄호의 조화로 오류 없이 완벽한 세계를 만들어 변수와 함수가 춤을 추고 알고리즘이 노래 를 불러

조코딩은 언어의 마법이야 Python, C++, Java, JavaScript 단어들이 춤을 춰 노래를 부르고 화면에 흘러 나오는 모습은 예술

실수와 버그에 좌절하기도 하지만 디버깅과 테스트로 문제를 해결해 조코딩은 결국 성취감의 나무 프로그래머의 마음을 높이 뛰게 만들어

세상을 변화시키는 조코딩 미래의 기술을 구현하는 키 끝없는 도전과 창의력의 발휘 프로그래머여, 너의 꿈을 향해 날아가라!





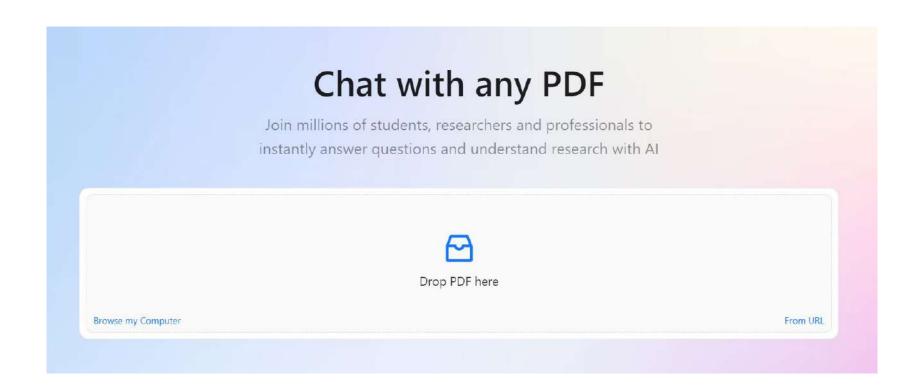
#### **ChatPDF**

JI LIMIA	PI KEY를 입력해주세요	
		•
PDF 파일	을올려주세요	
<b>①</b>	Drag and drop file here Limit 200MB per file • PDF	Browse files



#### 인공지능 시인 ChatPDF

#### **ChatPDF**



https://www.chatpdf.com/

### 기술 스택











**GPT** 



LLaMA 2 (오프라인)

# 개념설명

# 랭체인(LangChain)이란?

# <u>초거대 언어모델(LLM)</u> 어플리케이션 개발 <u></u>프레임워크

#### 초거대 언어 모델 (Large Language Model, LLM)?

1750억



**GPT** 

3400억



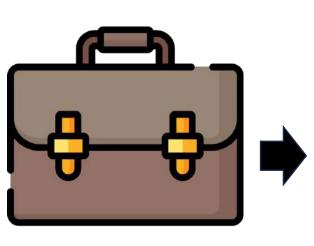
PaLM2

70억, 130억 등



**LLaMA** 

# 프레임워크? = 관련 도구 모음

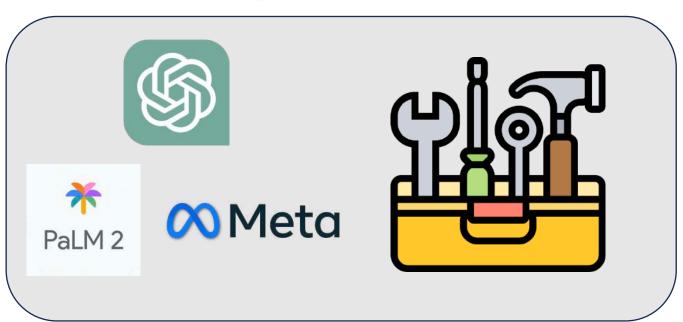


기계 수리용 프레임워크



# 랭체인(LangChain)이란?





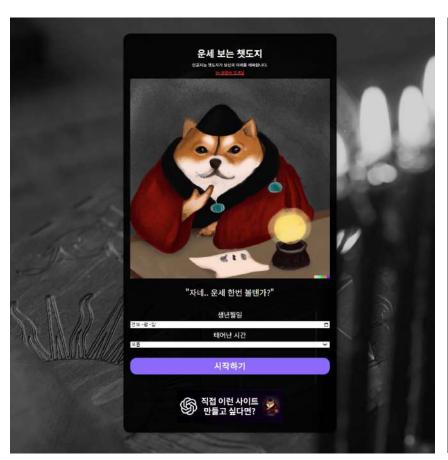
LLM 개발용 도구 모음

# 랭체인(LangChain)이란?



https://www.langchain.com

### 도구 모음을 왜씀?





⇒ 꼭 써야하는 건 아님 ㅇㅇ⇒ 따로 따로 갖다가 붙여도 됨

### 도구 모음을 왜씀?



여러가지 연결할때 편리함 교체 편리함 추상화로 코드 짧아짐

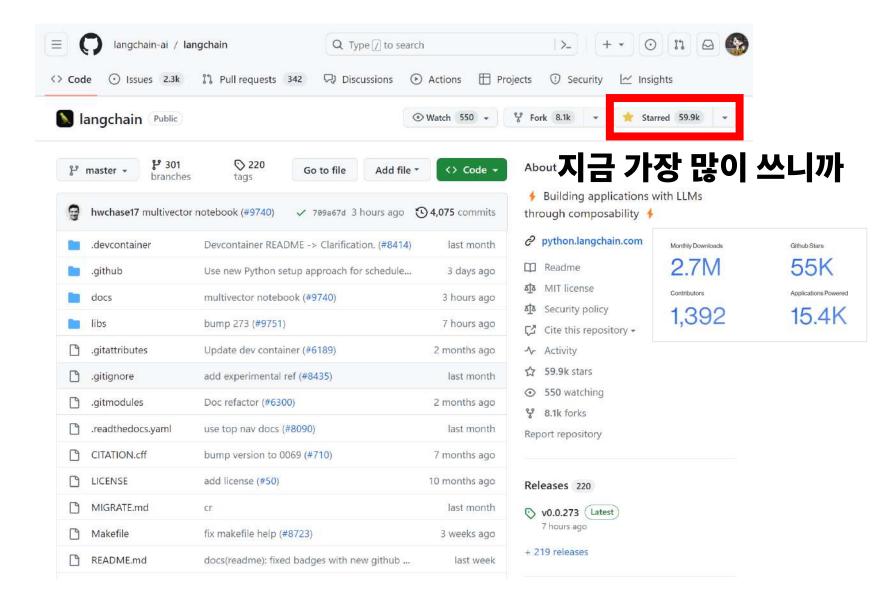
### 랭체인 말고는 없음?

⇒ 많음





### 근데 왜 랭체인 배움?



# Announcing our \$10M seed round led by Benchmark



It was only six months ago that we released the first version of LangChain, but it seems like several years. When we launched, generative AI was starting to go mainstream: stable diffusion had <u>just been released</u> and was captivating people's imagination and fueling an explosion in developer activity. Jasper appounded a <u>funding round</u>, and investors released the firs <u>Open DeepLearning.AI</u>

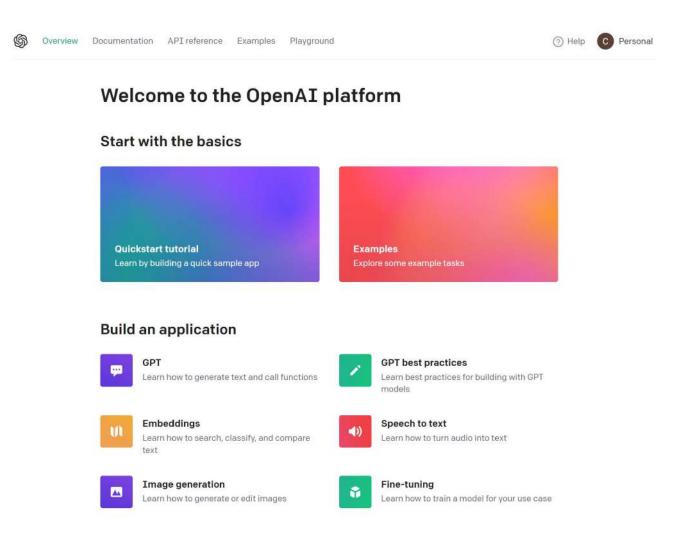




https://python.langchain.com/docs/get\_started/introduction.html

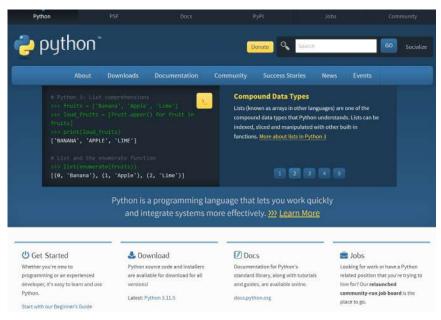
# 환경세팅

### OpenAl API Key 생성 & 카드 등록



https://platform.openai.com/

### Python 설치

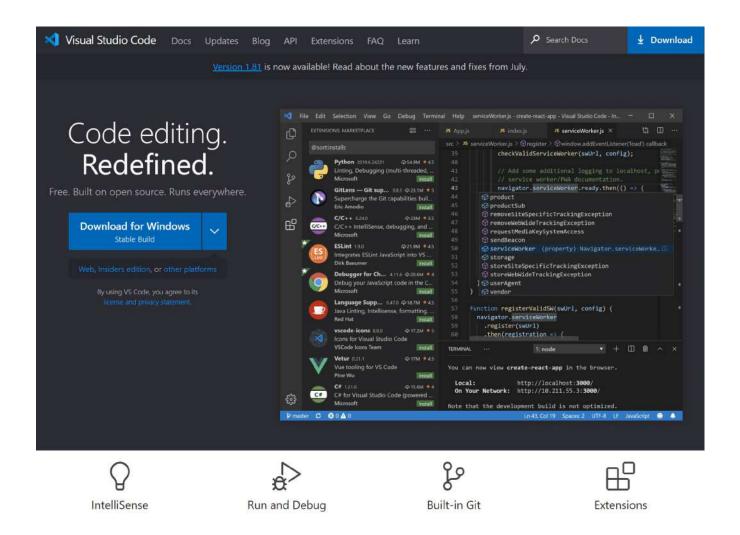


https://www.python.org/



https://www.youtube.com/playlist?list=PLU9-uwewPMe05-khW3YcDEaHMk\_qA-7ll

#### Visual Studio Code 설치

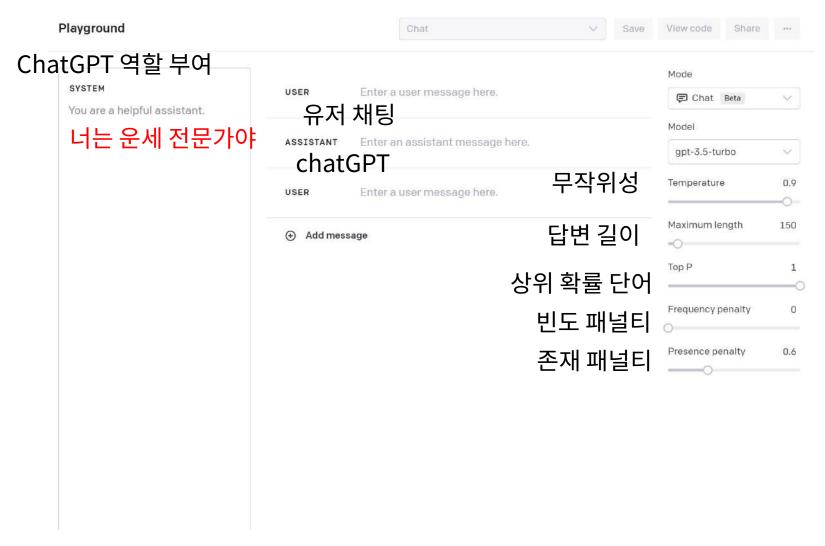


https://code.visualstudio.com/

# LangChain 시작하기

# LLM, Chat?

# Playground 체험



https://platform.openai.com/playground

#### **GPT**

내가 좋아하는 동물은 \_\_\_\_

강아지 0.5 고양이 0.3 늑대 0.1 아이폰 0.001

#### **ChatGPT**

User: 너가 좋아하는 동물이 뭐야?

Assistant: \_\_\_\_

# 인공지능 시인 만들기

#### Streamlit



Components

Generative Al Community

Blog

Sign in

Sign up

### A faster way to build and share data apps

Streamlit turns data scripts into shareable web apps in minutes. All in pure Python. No front-end experience required.

Try Streamlit now

Sign up for Community Cloud

https://streamlit.io/

#### **Streamlit Cloud**



#### **Community Cloud**

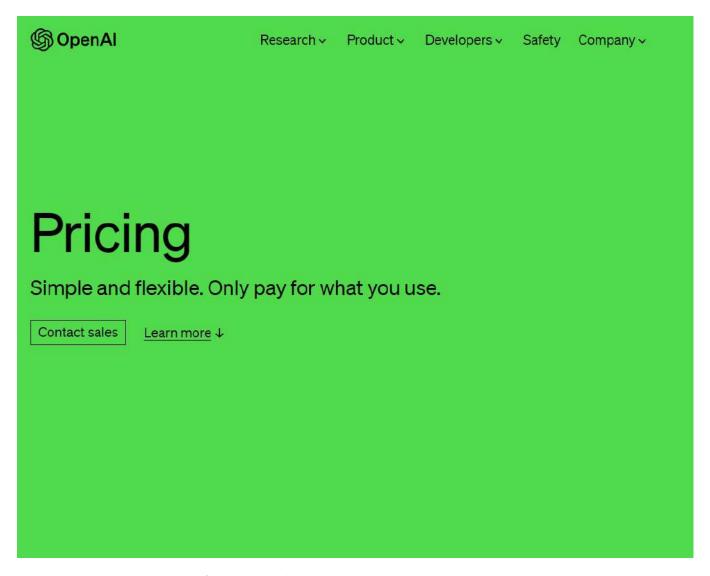
Deploy, manage, and share your apps with the world, directly from Streamlit — all for free.

Get Started

https://streamlit.io/cloud



### 가격



https://openai.com/pricing

# 돈 내기 싫다 회사 보안상 인터넷 안됨 정보 유출 우려됨

 $\bullet \bullet \bullet$ 

# 인터넷 없이 내 컴퓨터(로컬)에서만 LLM 돌리기

#### LLaMA 2

Meta Al Research Blog Resources About

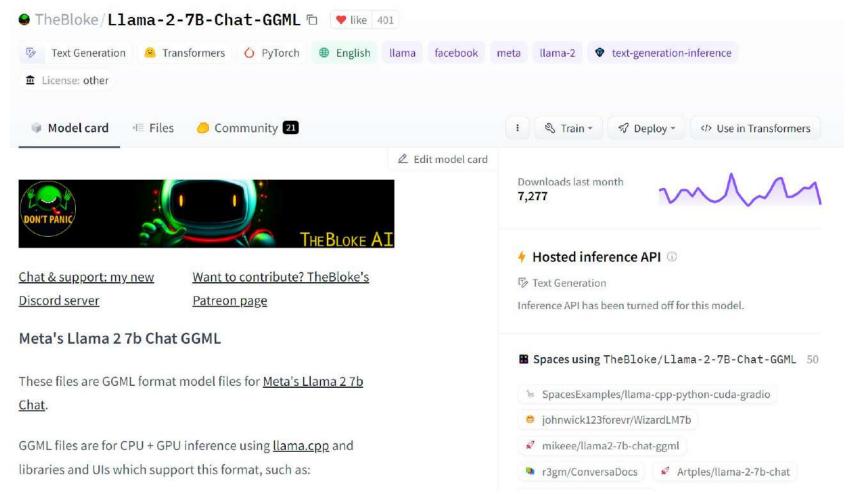
### Introducing Llama 2

The next generation of our open source large language model

Llama 2 is available for free for research and commercial use.

Download the Model

## LLaMA 2 경량화 버전

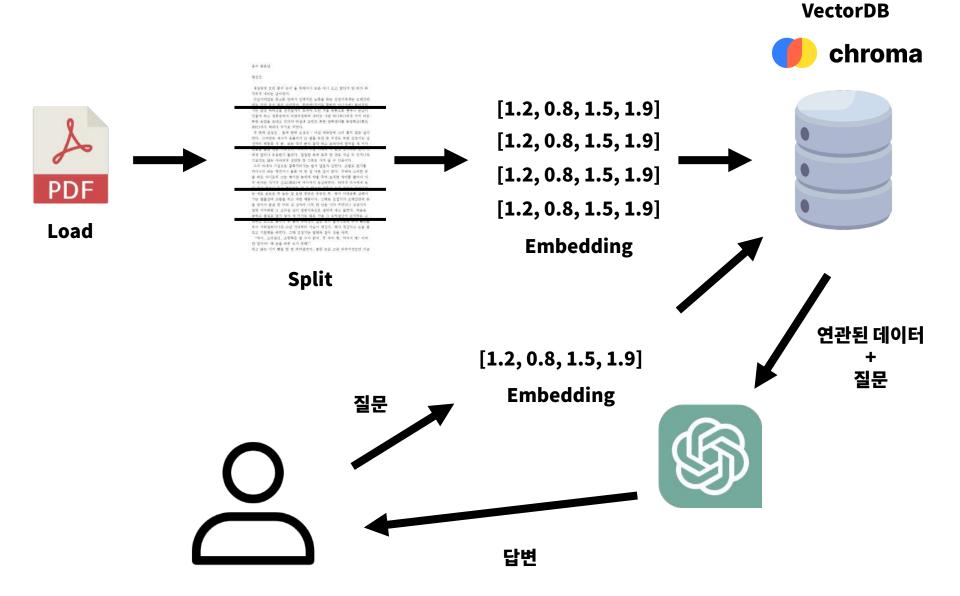


Georgi Gerganov + Machine Learning

https://huggingface.co/TheBloke/Llama-2-7B-Chat-GGML

## ChatPDF 만들기

## 전체 구조



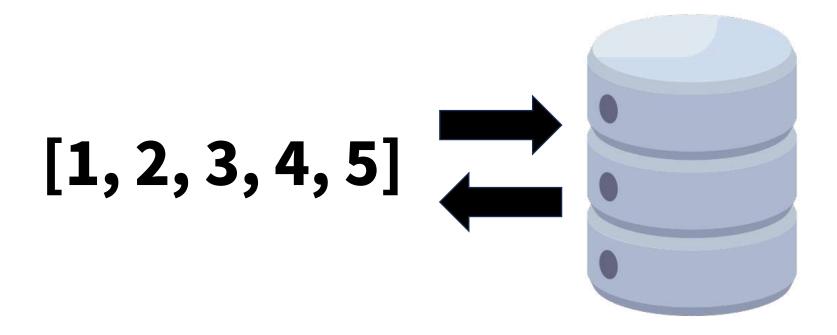
### 개념 설명

## VectorDB, 임베딩

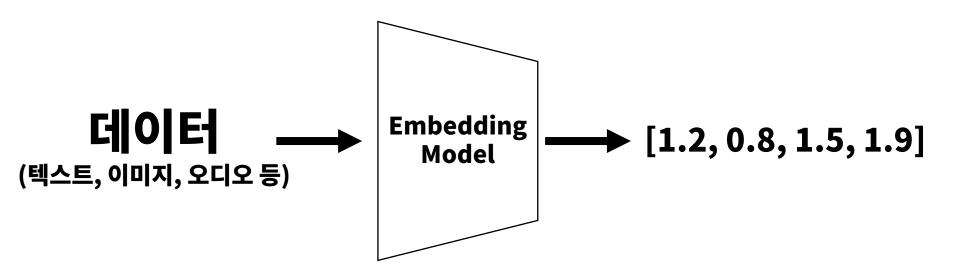
벡터(Vector) =

[1, 2, 3, 4, 5]

#### **VectorDB??**



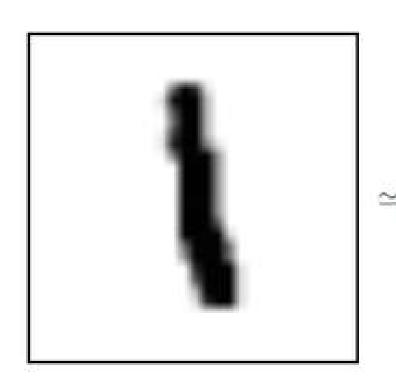
## 임베딩(Embedding)?



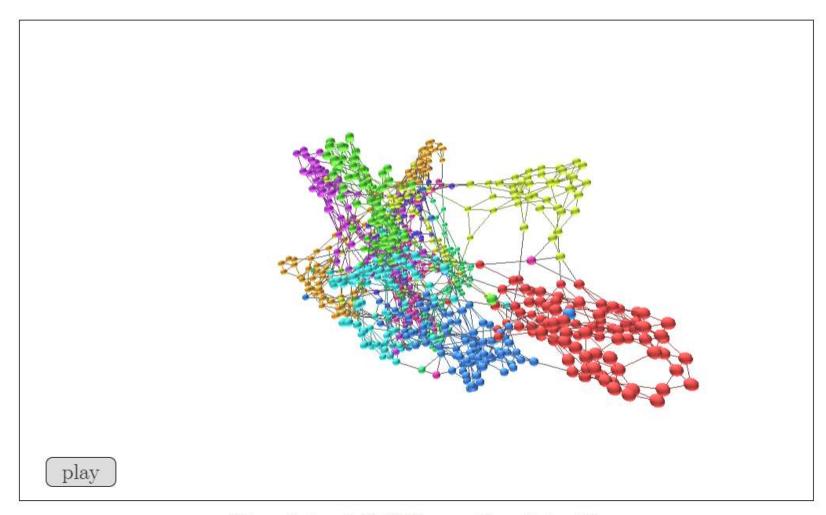
#### **MNEST**

1////// **3 3** 3 3 3 3 3 3 3 3 3 3 3 3 4 4 4 8 8 8 9 9 

### **MNEST**



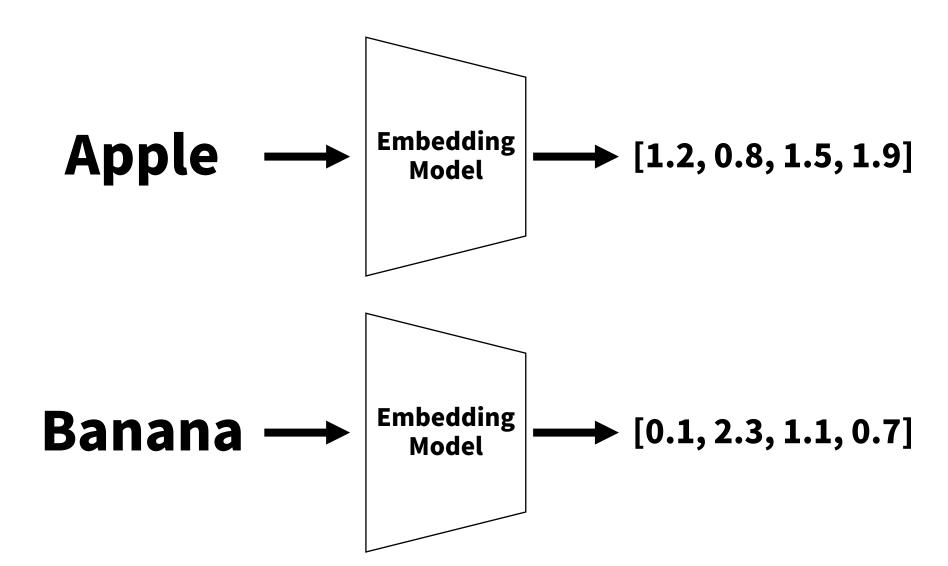
Ĭ	E.	0	W0	:54	ews	0	052	0.7	200	17.20	7007	000		
	000		0.1	:0		-0.0	.0	0.	0	0	0	00	0	0.0
	0	0	97	0	0	9	2	0	0	0	0	0	0	0
	.0	0	0	-0	0	0	-86		0	0.	40	0	0.	0
	.0	0	0.	0	0	0	17	100	0	0	0	0	0.	0
	9	0	9	0	0	0	T.	m	0	0	9	20	0	0
	0	0	0	0	0	0	15	(8)	[4]	0	9	U	0	0
	0	0	0	0	0	0	.0	000	128	0	0	-0	0.	0
	9	0	97	0	0	0	9	100	26	0	0	0	0	0
	0	0	0	0	0	0	0	31	130	0	9	. 0	0	0
	0	0	0	0	0	0	0	00	100	0	0	0	0	0
	97	0	93	0	0	0	9	110	100	[3]	9	90	0	0
	.0	0	0	0	0	0	0.		383	-1	0	.0	0	0
	0	0	0	0	0	0	0	0	0	0	0.	.0	0	0
	0	0	0	Ω	0	0	0	0	0	0	0	.0	0.	9

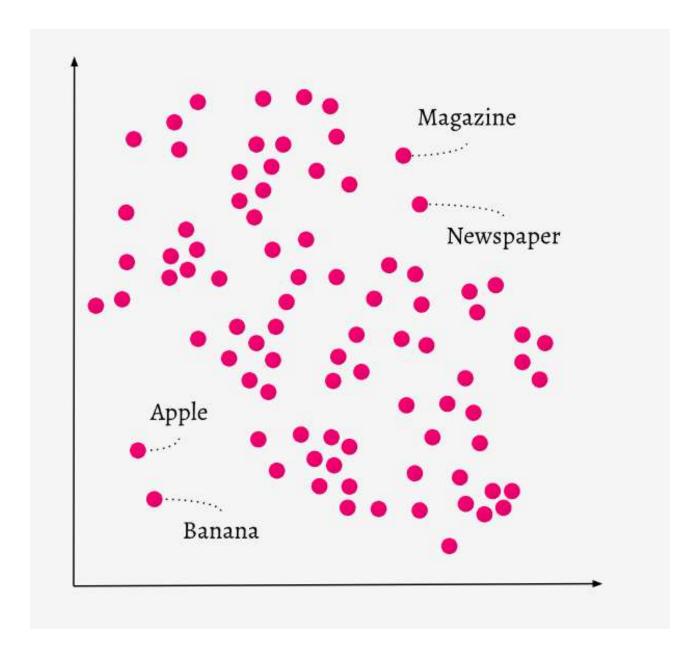


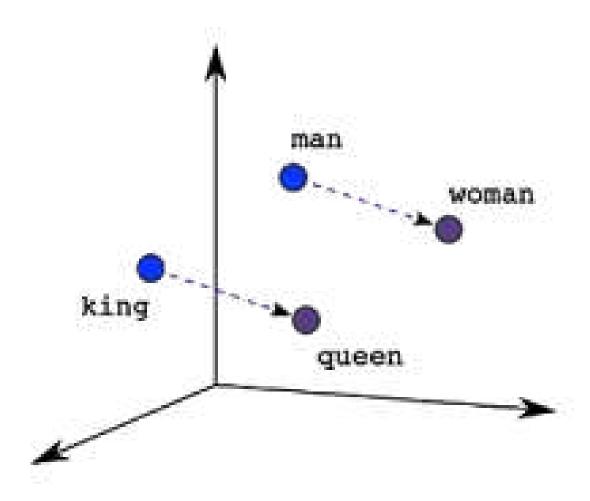
Visualizing MNIST as a Graph in 3D

(click and drag to rotate)

출처: https://colah.github.io/posts/2014-10-Visualizing-MNIST/







#### 꼬맨틀 - 단어 유사도 추측 게임



510번째 꼬맨틀의 정답 단어를 맞혀보세요.

정답 단어와 가장 유사한 단어의 유사도는 46.87 입니다. 10번째로 유사한 단어의 유사도는 39.54이고, 1,000번째로 유사한 단어의 유사도는 26.55 입니다.

추측할 단어를 입력하세요

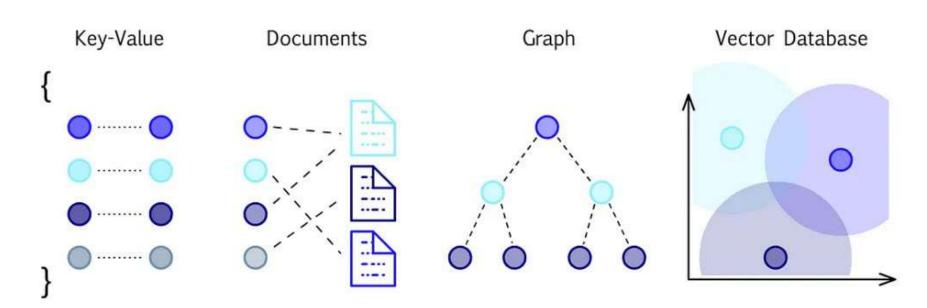
추측하기

포기하기



https://semantle-ko.newsjel.ly/

#### **VectorDB**



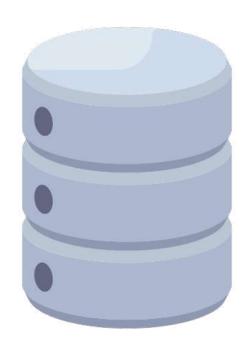
Source: Pinecone Docs

#### **VectorDB**









### 저작권 만료 소설

운수 좋은 날



https://gongu.copyright.or.kr/gongu/wrt/wrt/view.do?wrtSn=9002094&menuNo=200030

# 수익화

## 비용 🖡

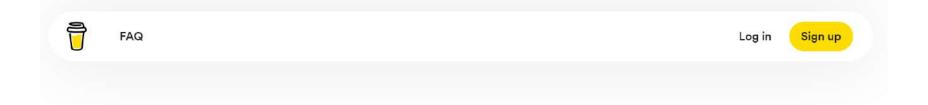
수익 1

#### **ChatPDF**





### Buy me a coffee



# A supporter is worth a thousand followers.

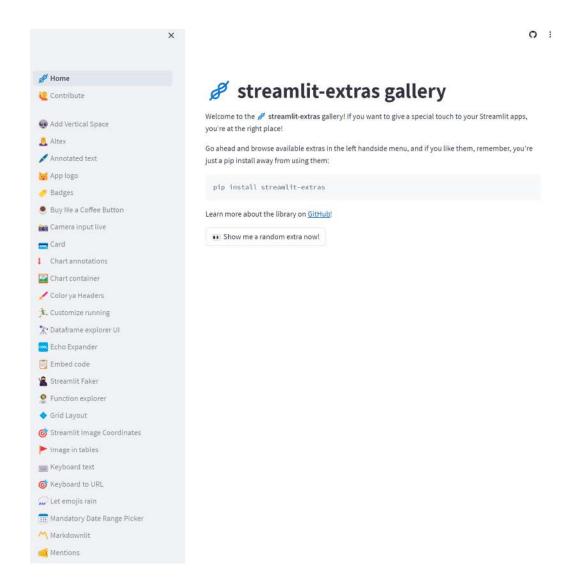
Accept donations. Start a membership. Sell anything you like. It's easier than you think.

buymeacoffee.com/yourname Start my page

It's free, and takes less than a minute.

https://www.buymeacoffee.com/

#### streamlit-extras



https://extras.streamlit.app/