

AI

NEWSLETTER

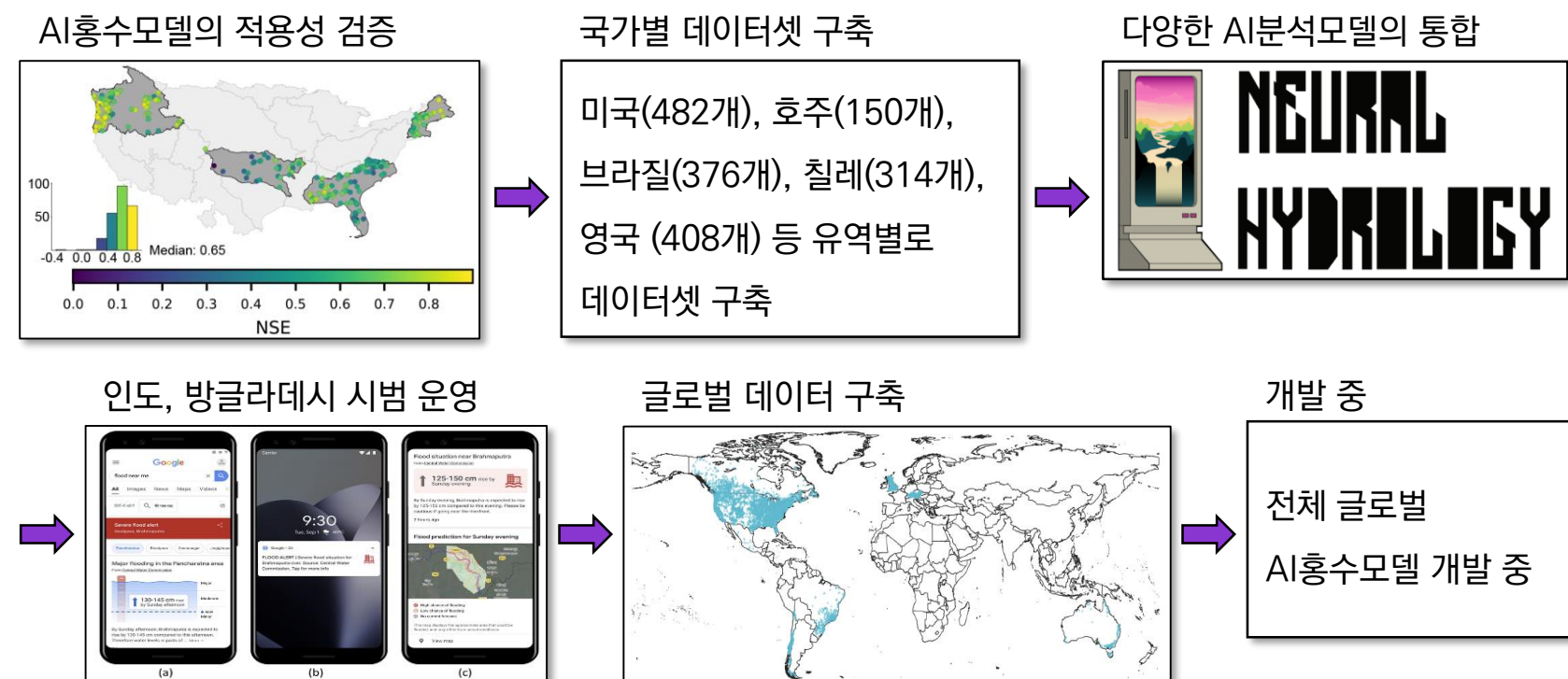
**K-water
AI Research Lab.**

31. July 2023

AI News

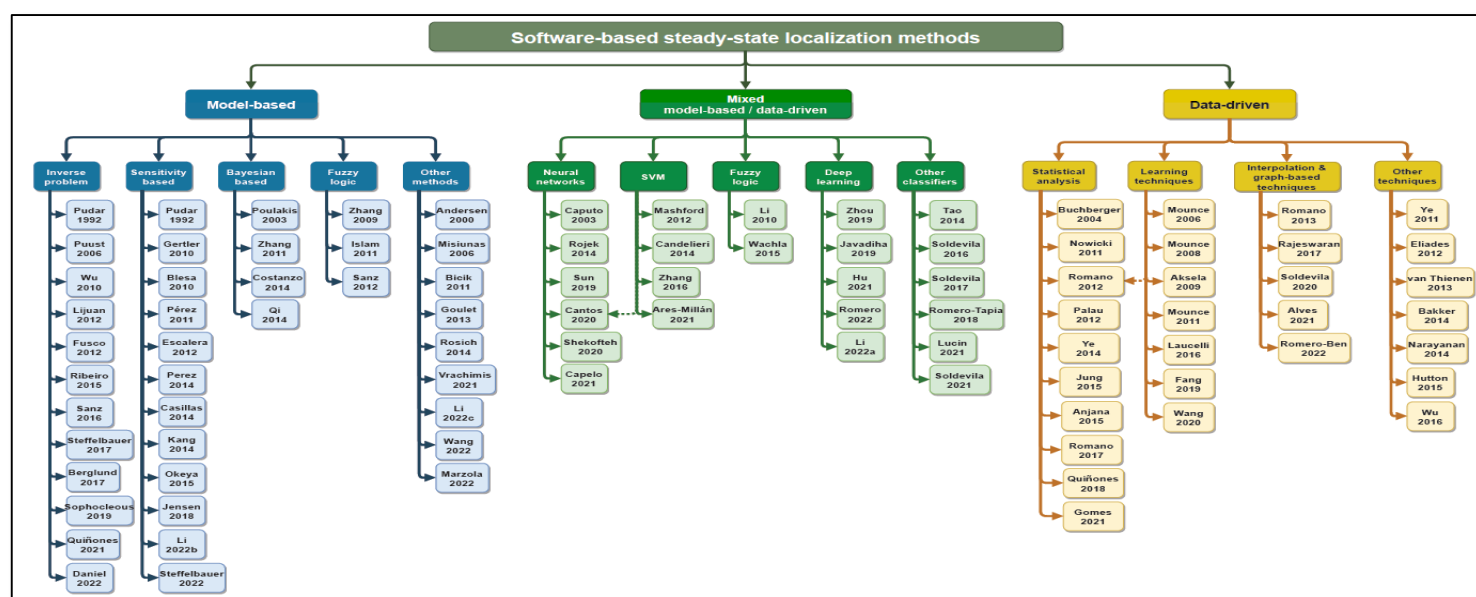
(수자원) 구글! 지구 홍수분석을 시로

구글은 세계 홍수분야 AI전문가들의 지원으로 글로벌 홍수분석이 가능한 시스템을 개발 중입니다.

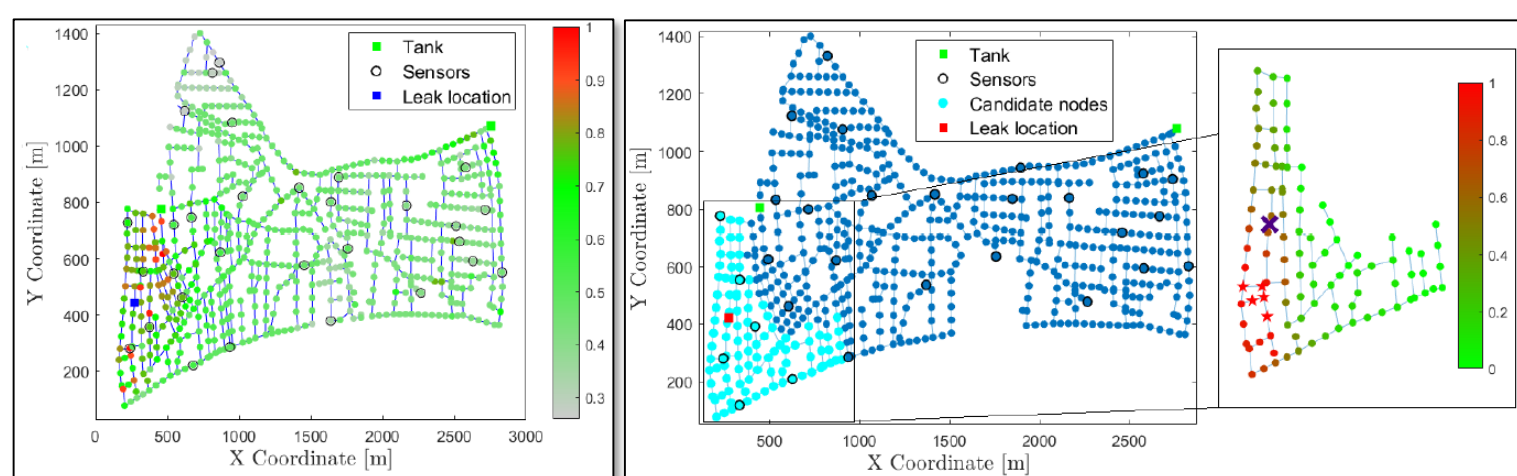


(수도) AI 누수 탐지

누수 탐지 관련하여 전통적인 방법 **Model-based**부터 AI 활용법 **Data-Driven**, 두 가지를 통합한 모델 **Mixed** 까지 다양한 방식으로 연구가 진행되고 있습니다.



대표적인 예로는 지중해 키프로스에서 수집된 데이터를 활용한 BattLeDIM 대회가 있습니다.



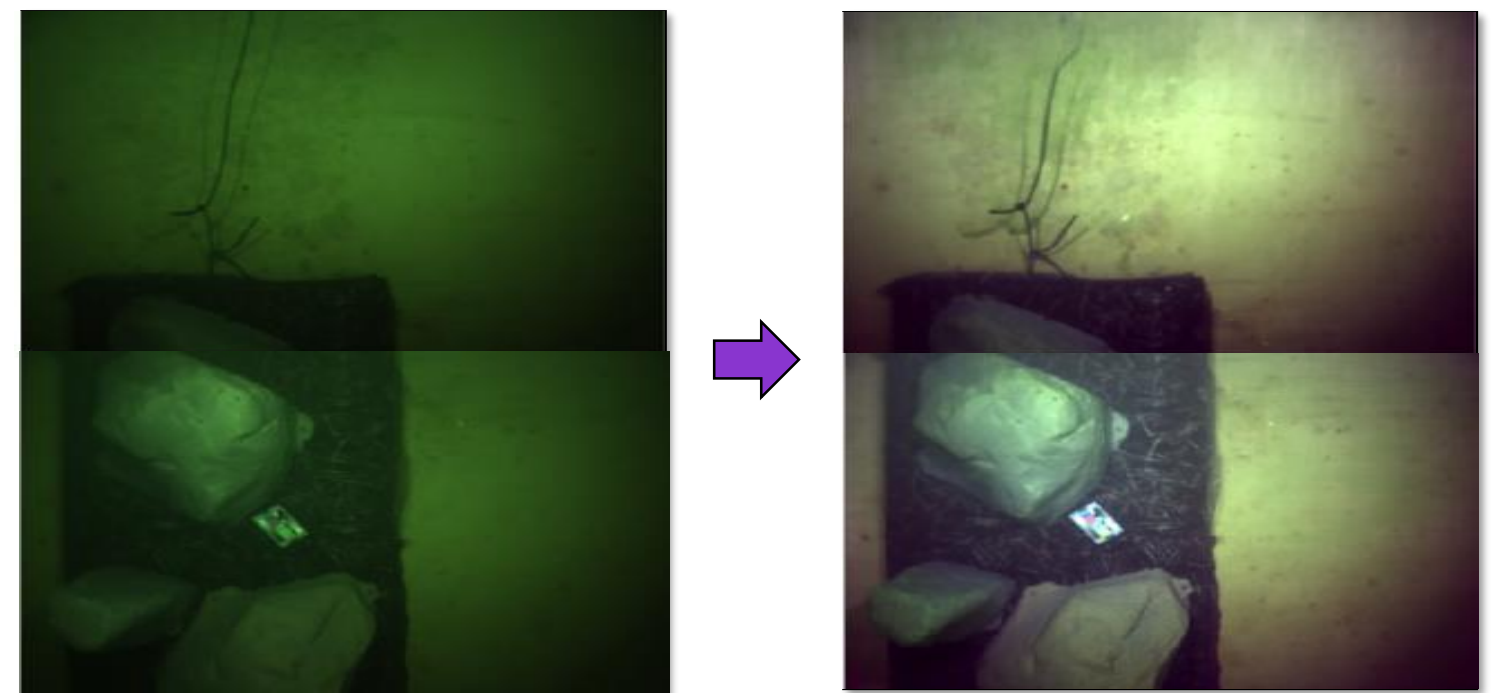
(1) 모델기반 분석

(2) Data-driven-candidate selection
(3) Data-driven and exact leak location

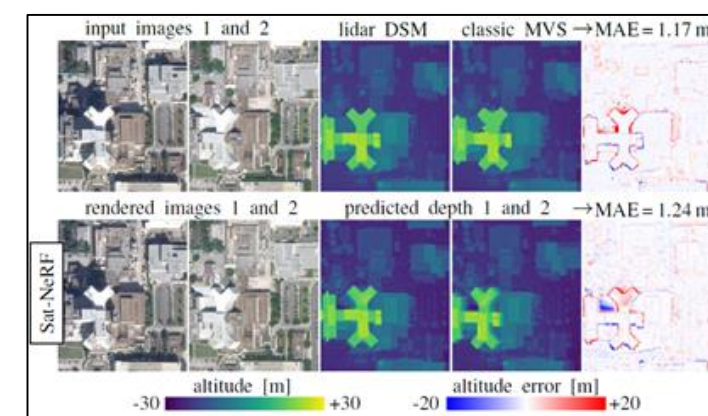
(NeRF) 사진으로 3D를??

사진만으로 3D를 만들 수 있는 기술이 있습니다. 바로 NeRF인데요, 넬프는 2020년 유럽학회 (ECCV)에서 발표되면서 많은 곳에서 사용하고 있습니다. 아래는 NeRF 기술을 이용해서 만든 3D입니다!! 정말 신기하지 않나요?

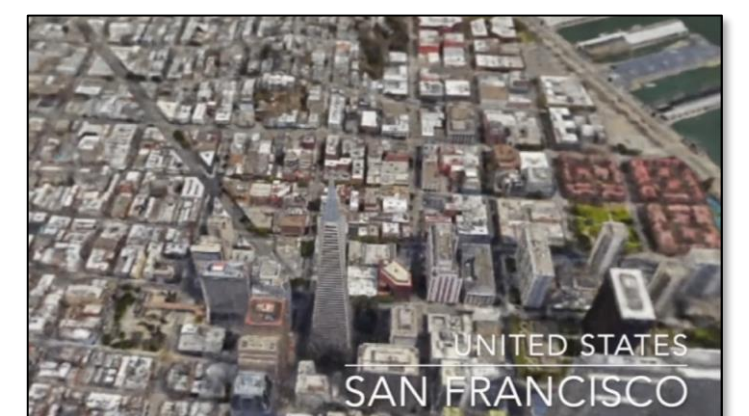
1. 수중 사진의 3D 변환



2. 위성영상 활용 3D 디지털화



3. 드론 활용 3D 디지털화



4. 3D영상에 홍수, 안개 등의 기상변화 적용 기술



사진을 이용해서 만든 3D



옆에 사진에 홍수 현상 표현

최근에는 영상 해상도를 높이고 빠른 학습을 위한 연구 및 GAN(Generative Adversarial Network) 기술 적용 연구도 진행되고 있습니다.



QR SCAN

Hands-On AI Project



QR SCAN

가독성을 높이기 위해 일부 세부 라인은 생략되어 있습니다. 전체 코드는 QR코드 링크를 참고해주세요.

나만의 ChatGPT 만들기

OpenAI 패키지를 사용해 나만의 ChatGPT를 만들어보겠습니다.

```
!pip install openai
```

사전학습 모델을 불러옵니다.

- 역할(role): 챗봇의 역할 (ex.helpful assistant)
- 온도(temperature): 답변의 창의성 정도

```
def generate_response(prompt, role, temperature):  
    response = openai.ChatCompletion.create(  
        model = engine, max_token = 1024,  
        message = [{ "role": "system", "content": f"You are a  
{role}." }, { "role": "user", "content": prompt }], )
```

사용할 모델(Engine)을 선택하고, 챗봇에게 역할(role)을 부여합니다.

```
engine_options = ["text-davinci-003", "text-davinci-002",  
"code-davinci-002", "gpt-3.5-turbo"]  
if not engine:  
    engine = "gpt-3.5-turbo"  
temperature_input = 0.7  
roles=["helpful assistant", "comedian", "medical doctor",  
"military personnel", "K-pop musician"]
```

마지막으로 Chatbot이 작동하는 코드를 작성합니다.

- 현재 선택 Engine = gpt-3.5-turbo
- Chatbot Role = K-pop musician
- Temperature = 0.7 (꽤나 창의적)

```
while True:  
    user_input = input("사용자: (*대화종료:Enter)")  
    if not user_input:  
        break  
    output = generate_response(user_input, selected_role,  
temperature_input)  
    print(f"챗봇: {output}")enter를 누르면 gpt-3.5-turbo]: ")
```

(결과 출력)

사용자 : 너의 대표곡은 뭐니?

챗봇 : 제 대표곡은 “불타오르네(FIRE)” 입니다. 이 곡은...(생략)...

사용자 : 대한민국 음악을 어떻게 생각해?

챗봇 : 저는 대한민국 음악을 매우 자랑스러워합니다. 대한민국은 글로벌 음악 씬에서 중요한 위치를 차지하고 있으며, ...(생략)...

ChatGPT 활용 데이터분석 (pandasAI)

Pandas... 파이썬 배우며 많이 들어보셨죠? 그렇지만 조작하는게 보통 어려운 일이 아닙니다. 요즘은 프롬프트만 넣으면 자동으로 답을 주는 PandasAI가 나왔습니다.

```
!pip install openai  
import pandas as pd  
from pandasai import PandasAI
```

샘플데이터를 연결해봅시다

```
df=pd.read_csv('/content/drive/MyDrive/ColabNotebooks/2023_July_Newsletter_5th/country.csv',encoding='CP949')
```

	country	total GDP	population	happiness_index					
0	US	19294482071552	331000000	6.94	4	Italy	1745433788416	60000000	6.38
1	UK	2891615567872	67000000	7.16	5	Spain	1181205135360	47000000	6.40
2	France	2411255037952	67000000	6.66	6	Canada	1607402389504	38000000	7.23
3	Germany	3435817336832	83000000	7.07	7	Australia	1490967855104	25000000	7.22
					8	Japan	4380756541440	126000000	5.87

거대언어모델(LLM)기반의 챗봇을 활용해서 질문을 입력 받고 대답하게끔 하는 구문입니다.

```
from pandasai.llm.openai import OpenAI  
llm = OpenAI(api_token="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")  
pandas_ai = PandasAI(llm, conversational=False)  
while True:  
    prompt = input("질문을 입력하세요. (*대화종료:Enter): ")  
    if not prompt:  
        break  
    output = pandas_ai(df, prompt=prompt, show_code=True)  
    print("Answer :", output)
```

아래 결과를 보면 잘 답변하는 걸 알 수 있습니다.

입력된 데이터는 다음과 같습니다: (중략)

질문을 입력하세요. (*대화종료:Enter): GDP를 기준으로 가장 부자인 나라와 가장 총GDP가 작은 나라는요?

Answer : 미국 스페인

질문을 입력하세요. (*대화종료:Enter): 두 번째로 행복한 나라와 5번째로 행복한 나라는요?

Answer : 호주 미국

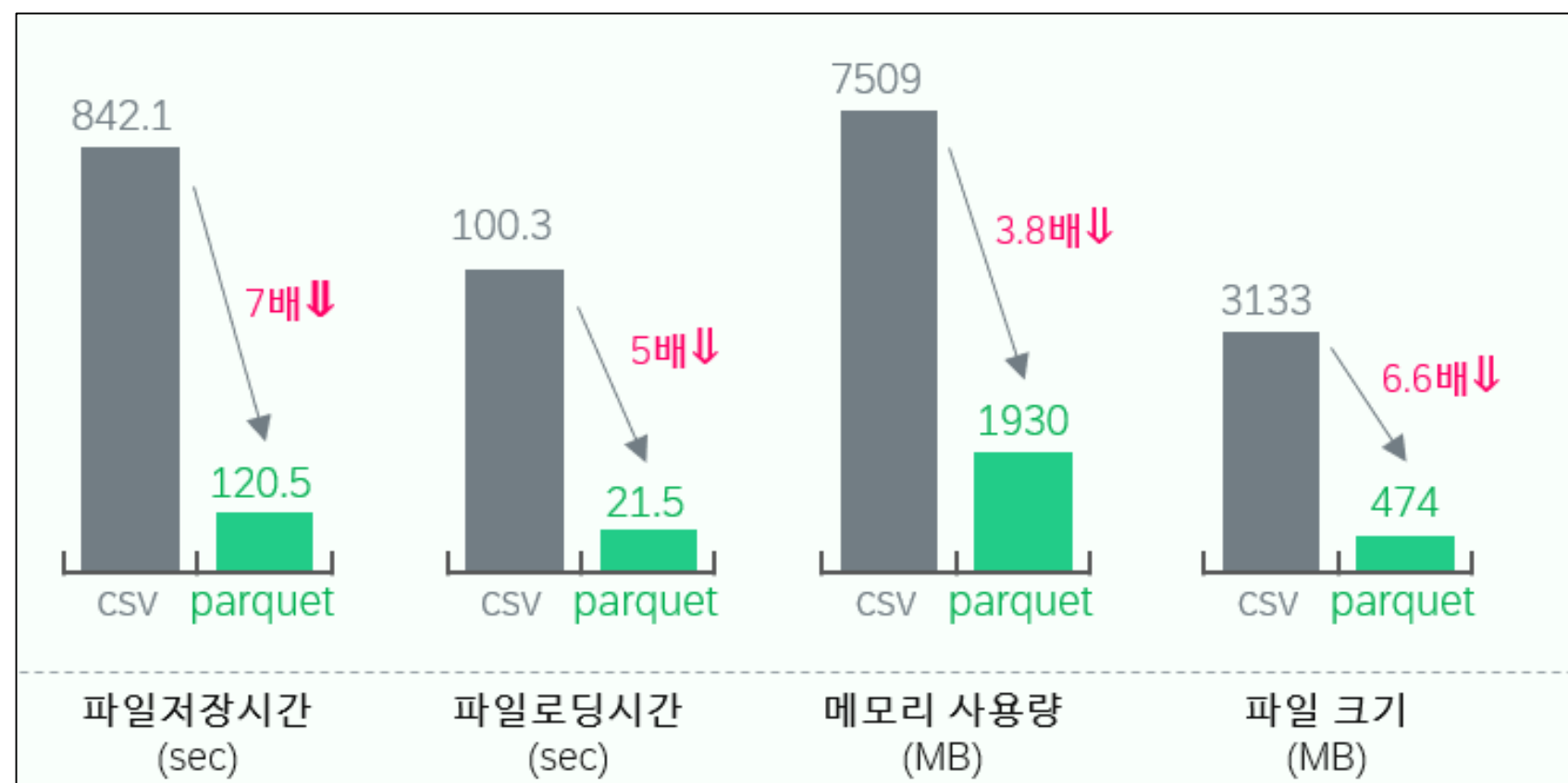
질문을 입력하세요. (*대화종료:Enter): 영국의 1인당 GDP는?

Answer : 1 43158.441312

TIPs

Parquet

Parquet은 csv와 같은 데이터 파일 형식입니다.
이 둘의 차이점은 Parquet의 데이터 압축률이 더
높고 데이터 처리를 더 빨리 할 수 있다는 점입니다.
대용량 데이터를 처리할 때 유용하겠죠?
아래는 성능 비교 표 입니다.



실제로 6,200만 건 정도의 대용량 데이터로 테스트
를 해봤습니다. CSV는 파일을 올리고 저장하는데
2분~9분정도 걸린 반면, Parquet은 1분도 안 걸렸
습니다. 파일 크기도 10배 정도 가볍습니다.
더욱이 오픈소스라 누구나 사용 가능합니다!

	Parquet	CSV	
read time	19초	2분 22초	7.4배 차이
write time	1분 1초	9분 46초	9.6배 차이
file size	0.5GB	5.75GB	11배 차이

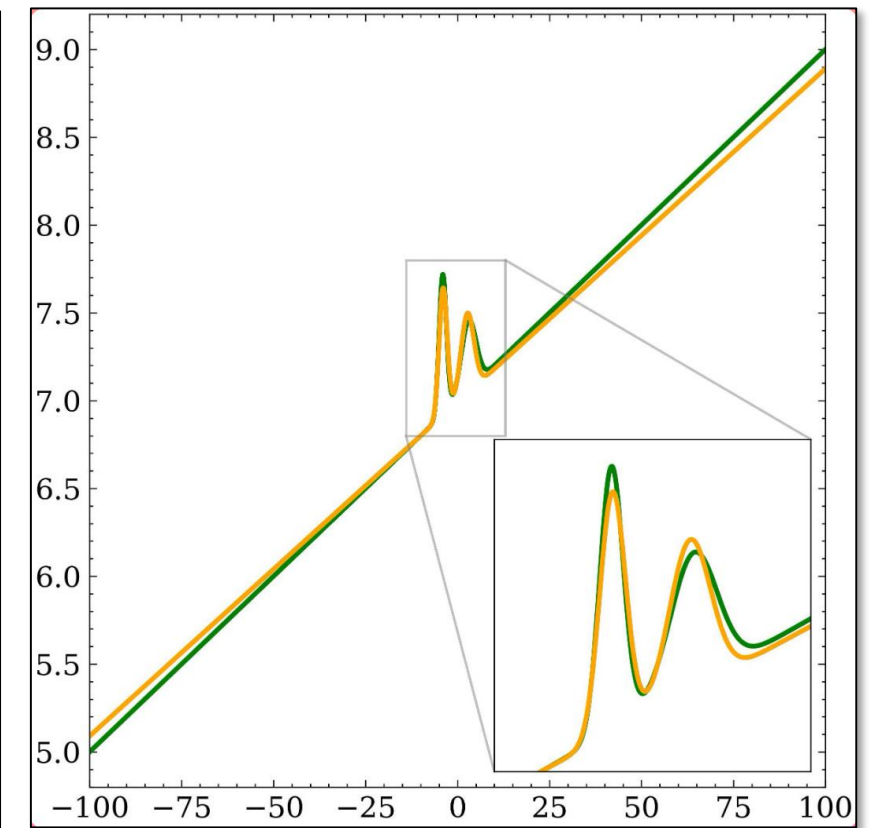
indicate_inset_zoom()

그래프에서 특히 강조하고 싶은 부분을 확대해서
보여주는 함수입니다.

```
#기본 plot 그리기
ax.plot(x,y)
axin = ax.inset(axes)
axin.plot(x,y)

#확대할 위치 선정
axin.set_xlim(-12,12)
axin.set_ylim(7,8)

#확대!!
ax.indicate_inset_zoom(axin)
```



PyGWalker

시각화의 최강자 태블로(Tableau)를 주피터에서 사용
할 수 있게 되었습니다! 설치도 간단하고 다양한 그
래프를 제공해 활용도가 높습니다.



K-water AI Lab.

국제대담회(ICOLD) 참석 (6/11-15)

- 15m 이상 되는 큰 댐들의 안정성을 검토하는데 활용될 K-water AI 기술 개발 및 비래비전 제시를 주제로 미국, 캐나다, 일본, 중국과 교류하는 시간을 가졌습니다.
- 특히, 한국대담회(K-water 주관)와 미국대담회(USSD), 캐나다대담회(CDA) 간 MOU를 체결했고, K-water의 첨단기술(드론영상기반 댐결함 검출, 사진기반 3D구축) 도입과 활용에 대해 발표했습니다.

2022년 분석평가 공로상 수상

MOU를 맺은 육군본부로부터 AI/빅데이터 분석 기법 및 데이터 공유를 통해 육군 분석평가 발전 도움에 대한 감사 공로상을 수상했습니다.



제3회 대국민 K-water AI경진대회

올해도 어김없이 AI경진대회가 열립니다!
아직 주제와 날짜는 미정이나 하반기에 진행될 예정이니 주제 추천과 많은 관심, 참여 부탁드립니다.

대한환경공학회 발표, 과제수행 워크숍

- 대한환경공학회(6.28)에서 “막여과 정수장 AI 적용 사례와 향후 비전”에 대해 발표했습니다.
- AI연구센터는 현재 12개의 협업(연구)과제와 8건의 컨설팅을 수행하고 있습니다. 워크숍을 통해 상호 연구 기술 공유와 효과적인 수행방안을 모색하는 시간을 가졌습니다.

AI 연구센터 세미나 개최

- AI연구센터는 매달 다양한 연구 주제로 세미나를 개최하고 있습니다.
- 지난 25일 개최한 “수문·기상분야와 AI에 대한 닥터컨퍼런스”에 많은 분들이 참석해주셨습니다.

날짜	발표자	내용
7.8	김세훈	수도미터 동파예측 AI모델 고도화 기술공유 세미나
7.25	신주영교수 윤성심박사	기상예측 분야에서 AI활용 연구 현황과 발전 방향 멀티모달 딥러닝의 수자원 분야 활용 연구



문의

주경원 선임 (7840)
이소령 사원 (7341)

깃허브 & 홈페이지

