

2025-2 졸업프로젝트

멀티모달 RAG와 성과 기반 분석을 활용한 식물 성장 분석 연구

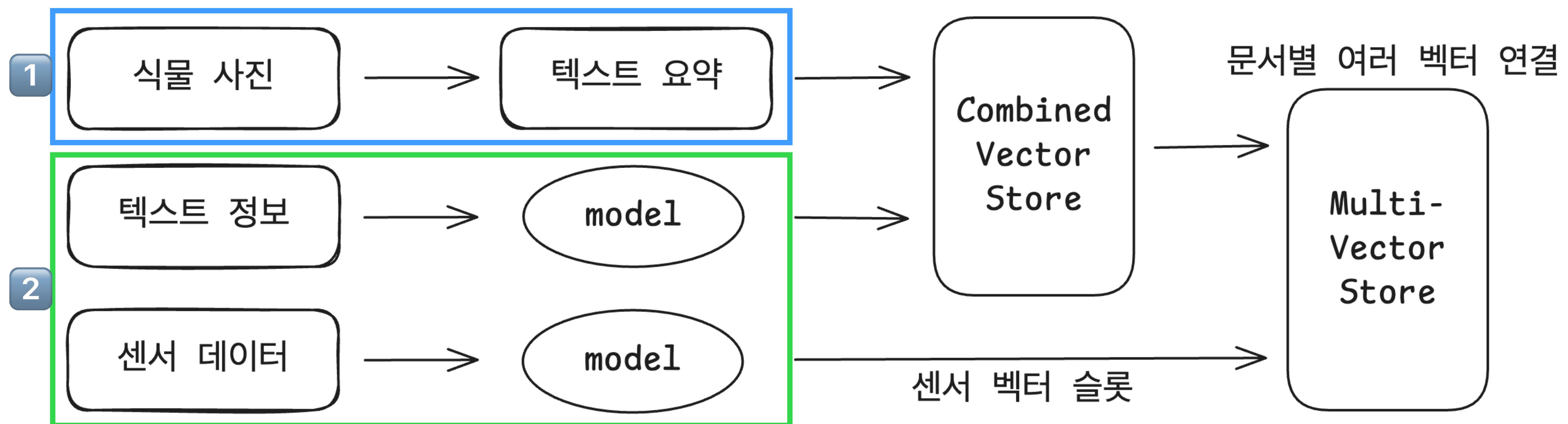
2025.10.01.

2021104417 우다현

앞으로 시도할 방법

앞선 두 연구의 방법을 참고하여 새로운 방식 설계

=> 텍스트+이미지 요약은 Combined Vector Store, 센서 데이터는 Multi-Vector 구조로 채택



1 기존 임베딩 모델 사용하기

- BLIP1 모델 사용

👉 Salesforce / **blip-image-captioning-base** 🔍 like 782

Follow 👉 Salesforce 1.79k

Image-to-Text Transformers PyTorch TensorFlow blip

image-captioning arxiv:2201.12086 License: bsd-3-clause

⋮ Train Deploy Use this model



추출된 텍스트 (캡션): "a green plant in a white pot on a white table"



추출된 텍스트 (캡션): "a green plant in a concrete pot"

1 기존 임베딩 모델 사용하기

프롬프트로 해결 시도



[목표: 잎_색상_상태] 프롬프트: "Describe the leaf color and texture. Is it wilted, dry, or yellow?"

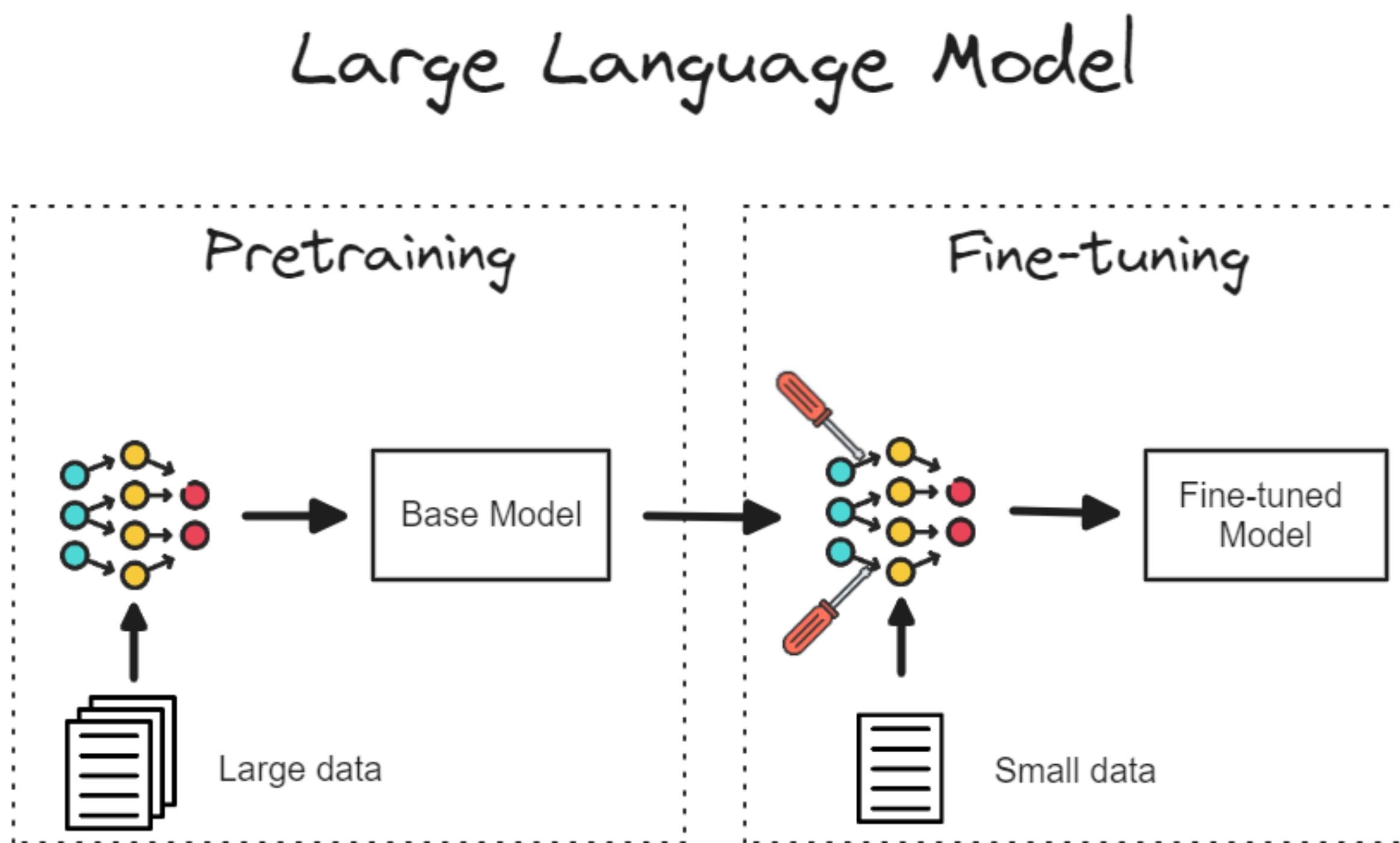
추출된 텍스트: "????????????"

[목표: 병징_및_피해] 프롬프트: "Identify any visible signs of disease, pests, or specific leaf spots/patches on the plant."

추출된 텍스트: "identify any visible signs of disease, pests, or specific leaf spots / patches on the plant.s / spots / spots on the plant"

1 파인 투닝 하기

- BLIP + [이미지 with 텍스트 데이터]
- 존재하는 이미지 + 텍스트 매칭하여 학습 예정



1 Gemini API 사용하기

이미지 학습에 과도한 리소스 사용시 Gemini API 사용예정

[프롬프트] :

"이 이미지 속 **식물의 건강 상태만**을 설명하는 문장을 **하나** 생성하세요.

배경 정보는 모두 제외하고, 잎의 **색상 변화**, **처짐 여부**, **전반적인 활력**에 대해서만 상세히 서술하세요.



[캡션] :

잎이 심하게 황변(노란색으로 변색)되어 있고, 아래쪽 잎들은 갈색으로 변하며 축 늘어져 처진 상태이다.

잎의 활력과 긴장도가 매우 낮아 심각한 스트레스를 겪고 있는 식물이다.

2 멀티 벡터

이미지 캡션 작업 전 멀티 벡터 임베딩

텍스트 데이터 → 벡터 변환

- 모델: **BGE-m3-ko** (한국어 특화 임베딩 모델)
- 처리 과정:
 1. JSON 파일에서 식물 상태 설명을 문장으로 생성
(예: “선인장은 베란다에서 자라는 건생식물이며, 환경은 건조 상태이고 …”)
 2. HuggingFace 임베딩 모델을 통해 **1024차원 벡터**로 변환
- 활용: 자연어 질의(“물이 적게 필요한 식물”)와의 의미적 유사도 검색

2 멀티 벡터

이미지 캡션 작업 전 멀티 벡터 임베딩

센서 데이터 → 벡터 변환

- 모델: **TabNet** (표 형식 데이터 전용 딥러닝 모델)
- 처리 과정:
 1. 센서 값을 배열 형태로 정리
(예: [기온, 습도, CO₂, 토양습도, pH ...])
 2. TabNet을 학습시킨 뒤, **latent embedding layer**에서 센서 벡터 추출
→ 보통 **64차원 벡터**
- 활용: 특정 센서 패턴과 유사한 식물 상태 검색

2 멀티 벡터

이미지 캡션 작업 전 멀티 벡터 임베딩

멀티벡터 통합

- 두 종류의 벡터를 **같은 문서 단위(id)**에 저장
 - text_vector (1024차원)
 - sensor_vector (64차원)
- 검색 시:
 - 텍스트 쿼리 → text_vector와 비교
 - 센서 쿼리 → sensor_vector와 비교
 - 두 점수를 가중합(a-블렌딩)하여 최종 결과 도출 (**Late Fusion** 방식)

감사합니다