

제시하신 YAML 형식의 코딩 그룹을 한국어로 해석해 드리겠습니다. 이 코드는 점수 산출 및 프로파일링 엔진의 설정 파일로 보이며, 학생 또는 대상자의 여러 지표를 계산하고 특정 조건에 따라 프로파일을 분류하며 맞춤형 권고 사항을 생성하는 데 사용되는 규칙을 정의하고 있습니다.

❖ 주요 구성 요소 및 해석

1. version

해석: 버전

설명: 설정 파일의 버전 정보입니다. (예: )

2. norms (규준/지표)

설명: 다양한 출처의 데이터를 기반으로 점수를 계산하고 정규화하는 방법을 정의합니다.

구성:

id: multicultural_adaptation (다문화 적응 지표)

source:multicultural_form (다문화 설문지)

compute:sum(items[Q1:Q5]) (Q1부터 Q5까지 항목들의 합계로 계산)

normalize:minmax(0, 100) (최소값 0, 최대값 100 범위로 정규화)

id: attention_index (주의 집중 지표)

source:neuroharmony (신경 조화/뇌파 관련 데이터 출처로 추정)

compute:weighted_sum({beta: 0.4, theta: 0.3, alpha: 0.3}) (베타파, 세타파, 알파파 등의 가중 합계로 계산)

normalize:minmax(0, 100) (최소값 0, 최대값 100 범위로 정규화)

id: study_habit (학습 습관)

source:school_form (학교 설문지)

compute:sum(items[H1:H4]) (H1부터 H4까지 항목들의 합계로 계산)

normalize:minmax(0, 100) (최소값 0, 최대값 100 범위로 정규화)

3. profiles (프로파일)

설명: 계산된 지표 값(norms)을 기반으로 특정 조건을 만족할 때 '리스크 플래그'와 같은 프로파일을 할당하는 규칙을 정의합니다.

구성:

id: risk_flags (위험/리스크 플래그)

rules (규칙):

규칙 1:

when:attention_index < 35 and multicultural_adaptation > 70 (주의 집중 지표가 35 미만이고 다문화 적응 지표가 70 초과일 때)

then:

label:집중/문화 적응 동시 리스크

severity:high (심각도: 높음)

규칙 2:

when:study_habit < 40 (학습 습관이 40 미만일 때)

then:

label:학습 습관 개선 필요

severity:medium (심각도: 보통)

4. recommendations (권고 사항)

설명: 특정 ID를 가진 권고 사항을 생성하는 방법을 정의합니다. 여기서는 RAG(Retrieval-Augmented Generation) 모델을 사용하도록 설정되어 있습니다.

구성:

id: four_week_plan (4주 계획)

use_rag:true (RAG 모델 사용)

prompt (프롬프트):

내용: "아래 프로필을 바탕으로 4주 개입 계획을 작성하세요."

필수 포함 정보:

학년/연령

리스크 플래그 (risk_flags)

다문화 적응 지표 (multicultural_adaptation)

보호자 협력 방안

출력 형식: 요약, 주차별 목표, 가정/학교 실천, 평가 기준

요약하자면, 이 코드는 대상자의 다문화 적응, 주의 집중, 학습 습관을 정량화하고, 이 점수

를 바탕으로 위험 요소를 식별하여()¹, 최종적으로 개인 맞춤형 4주 개입 계획을 인공지

능이 생성()²)하도록 지시하는 설정 파일입니다.

제시하신 Python 코드는 LangGraph 기반 파이프라인의 상태(State) 정의를 담고 있으며, 데이터 처리 과정 전반에 걸쳐 정보를 추적하고 관리하는 데 사용됩니다.

❖ graph/state.py 한국어 해석 (LangGraph 상태 정의)

이 코드는 Python의 Pydantic 라이브러리를 사용하여 데이터 모델을 정의합니다. Pydantic 모델은 데이터 유효성 검사 및 설정 관리에 유용합니다.

1. Student 클래스 (학생 정보)

설명: 처리 대상인 학생의 기본 정보를 담는 모델입니다.

구성 요소:

id (str): 고유 식별자 (필수)

name (Optional[str] = None): 이름

grade (Optional[str] = None): 학년

school (Optional[str] = None): 학교

guardians (Optional[List[str]] = None): 보호자 연락처 또는 이름 목록

2. PipelineState 클래스 (파이프라인 상태)

설명: 데이터 분석 및 보고서 생성 파이프라인의 전체 작업 흐름 상태를 관리하는 핵심 모델입니다. 각 단계의 결과가 여기에 저장됩니다.

구성 요소:

student (Student): 위에서 정의된 학생 정보 객체.

raw_inputs (Dict[str, Any]): 처리 전의 원본 데이터 (파일이나 텍스트).

예시: forms_csv (설문지 CSV), neuro_ban (신경/뇌파 측정 데이터), neuro_pdf_text (신경 관련 PDF 텍스트)

validated (Dict[str, Any]): 유효성 검증 결과.

예시: schema_ok (스키마 적합 여부), anomalies (이상치/결측치 정보)

scores (Dict[str, Any]): 점수 계산 결과.

예시: scales (척도 점수), flags (위험 플래그/표식, 앞선 YAML 파일의 결과)

analysis (Dict[str, Any]): 분석 및 해석 결과.

예시: summaries (요약), insights (통찰/핵심 발견)

report (Dict[str, Any]): 최종 보고서 정보.

예시: md (마크다운 텍스트), pdf_path (PDF 파일 경로), canva_link (캔바 디자인 링크)

log (List[Dict[str, Any]]): 파이프라인에서 발생하는 주요 이벤트 기록 목록 (디버깅 및 감사에 사용).

예시: events (발생한 단계/이벤트), versions (사용된 모델 버전)

`log_event` (메서드): 로그 항목을 추가하는 함수.

요약: 이 코드는 학생을 중심으로, 원시 입력부터 데이터 검증, 점수 산출, 분석, 최종 보고서까지의 모든 처리 단계를 체계적으로 저장하고 관리하는 데이터 컨테이너 역할을 수행합니다. 이는 LangGraph(상태 기반 그래프 실행 프레임워크)에서 상태를 노드 간에 전달하고 업데이트하는 데 필수적입니다.

제시하신 Python 코드는 LangGraph 파이프라인의 첫 번째 노드 중 하나로, 외부에서 제공되는 입력 데이터를 수집하고 파이프라인 상태에 저장하는 역할을 합니다.

graph/nodes/ingest_inputs.py 한국어 해석 (입력 수집/저장)

1. 전역 설정 및 준비 (Global Setup)

설명: 필요한 모듈을 불러오고, 원시 데이터를 저장할 디렉토리를 설정합니다.

DATA_RAW 변수는 현재 스크립트 파일 위치에서 상위 디렉토리를 거슬러 올라가서 찾은 data/raw 경로를 지정합니다.

`DATA_RAW.mkdir(parents=True, exist_ok=True)`: 이 경로에 디렉토리가 없으면 생성하고, 이미 있어도 오류를 발생시키지 않도록 합니다.

2. run 함수 (노드 실행)

역할: 이 함수는 노드의 핵심 로직을 담고 있으며, LangGraph에서 이 노드가 실행될 때 호출됩니다.

입력:

`state: PipelineState`: 현재 파이프라인의 전체 상태 객체 (이전 요청에서 해석했던 PipelineState).

`payload: Dict[str, Any]`: 이 노드로 전달되는 새로운 입력 데이터 (예: 파일 경로).

주요 로직:

`입력 저장: state.raw_inputs.update(payload)`

payload에 담겨 온 모든 키-값 쌍(예: forms_csv, neuro_pdf 파일 경로)을 PipelineState 객체의 raw_inputs 덕셔너리에 추가하거나 업데이트합니다. 이로써 파이프라인의 다음 노드들이 이 원시 입력 정보에 접근할 수 있게 됩니다.

이벤트 기록: state.log_event("ingest_inputs", {"keys": list(payload.keys())})

상태 객체의 log 목록에 이벤트를 기록합니다. 이벤트 이름은 "ingest_inputs"이며, 어떤 키의 데이터가 수집되었는지(payload의 키 목록)를 기록합니다.

상태 반환: return state

업데이트된 PipelineState 객체를 다음 노드로 전달하기 위해 반환합니다.

요약: 이 코드는 LangGraph 파이프라인에 초기 데이터를 주입하는 역할을 합니다. 외부에서 제공된 원시 입력 정보(주로 파일 경로)를 받아 PipelineState의 raw_inputs 필드에 저장하고, 이 작업을 로그에 기록한 후 업데이트된 상태를 다음 단계로 넘깁니다.

제시하신 Python 코드는 LangGraph 파이프라인의 노드 중 하나로, 수집된 원본 데이터의 유효성을 검증하는 역할을 합니다. 특히, 설문지 파일에 필수 컬럼(열)이 모두 포함되어 있는지 확인하고 결측치(누락된 컬럼) 정보를 상태에 기록합니다.

 graph/nodes/validate_schema.py 한국어 해석
(Pydantic 검증/결측치 확인)

1. 전역 설정 (Global Settings)

REQUIRED_COLUMNS: 데이터 파일에 반드시 존재해야 하는 컬럼들의 목록을 정의합니다.

예시): 학생 ID(student_id)와 설문 항목 Q1부터 Q5까지.

2. run 함수 (노드 실행)

역할: PipelineState에 저장된 입력 파일(설문지 CSV)을 읽어 정해진 스키마(형식)에 맞는지 확인합니다.

입력: state: PipelineState (현재 파이프라인의 전체 상태 객체).

주요 로직:

입력 파일 경로 가져오기:

forms_csv = state.raw_inputs.get("forms_csv")를 통해 원시 입력(raw_inputs)에 저장된 설문지 CSV 파일 경로를 가져옵니다.

검증 변수 초기화:

```
schema_ok, anomalies = True, [] : 기본적으로 스키마는 **정상(True)**으로 설정하고,  
이상 징후 목록(anomalies)은 비워둡니다.
```

CSV 파일 처리 및 검증:

```
if forms_csv:: 파일 경로가 존재할 경우에만 처리합니다.
```

```
df = pd.read_csv(forms_csv): forms_csv 경로의 파일을 Pandas DataFrame으로 읽어옵니다.
```

누락된 컬럼 확인:

```
missing = [c for c in REQUIRED_COLUMNS if c not in df.columns] :  
REQUIRED_COLUMNS 목록을 순회하며 DataFrame에 포함되어 있지 않은 컬럼들을 찾아  
missing 목록에 저장합니다.
```

검증 결과 업데이트:

```
if missing:: 누락된 컬럼이 하나라도 발견되면,
```

```
schema_ok = False: 스키마 적합 여부를 **비정상(False)**으로 변경하고,
```

```
anomalies.append(...): anomalies 목록에 어떤 컬럼이 누락되었는지 기록합니다.
```

상태 저장 및 로깅:

```
state.validated = {"schema_ok": schema_ok, "anomalies": anomalies}: 검증 결과를  
PipelineState의 validated 필드에 저장합니다.
```

```
state.log_event("validate_schema", state.validated): 검증 결과와 함께 이벤트를 로그에 기  
록합니다.
```

```
상태 반환: return state (업데이트된 상태를 다음 노드로 전달).
```

요약: 이 코드는 입력 데이터의 무결성을 확인하는 단계로, 설문지 CSV 파일이 후속 처리 단계에서 요구하는 **필수 데이터 열(스키마)**을 모두 갖추고 있는지 검사하여, 그 결과를 PipelineState의 **validated** 섹션에 저장합니다. 이는 데이터 품질 문제를 조기에 식별하여 오류를 방지하는 데 중요합니다.

제시하신 Python 코드는 LangGraph 파이프라인의 핵심 노드 중 하나인 **점수 산출 엔진**의 역할을 합니다. 이는 이전에 해석했던 score_engine.yaml 설정 파일을 로드하여 지표를 계산하고, 그 결과를 바탕으로 리스크 플래그를 할당하는 기능을 **시연(데모)**하는 코드입니다.

 graph/nodes/score_engine.py 한국어 해석
(YAML 규칙 로더 + 간단 계산)

1. 설정 파일 로드 (Configuration Loading)

경로 설정:

CONFIG 변수는 현재 스크립트 위치를 기준으로 상위 디렉터리를 거슬러 올라가 config(score_engine.yaml) 파일을 가리킵니다. 이 YAML 파일은 점수 계산 및 리스크 규칙을 정의합니다.

YAML 로드:

with open(CONFIG, "r", ...): score_engine.yaml 파일을 읽어 yaml.safe_load(f) 함수를 사용하여 Python 객체(cfg)로 변환합니다. (이후 코드에서는 cfg를 사용하지 않아 예제의 단순함을 보여줍니다. 실제 시스템에서는 이 cfg를 기반으로 복잡한 계산을 수행합니다.)

2. 정규화 도우미 함수 (Normalization Helper)

_minmax(x, xmin, xmax):

역할: 특정 값  를 최소값 \$x_{\min}\$과 최대값 \$x_{\max}\$를 기준으로 에서  사이의 범위로 정규화하는 함수입니다. (Min-Max Scaling)

수식 해석: 

3. run 함수 (노드 실행)

입력 처리:

forms_csv: PipelineState의 raw_inputs에서 설문지 파일 경로를 가져옵니다.

df = pd.read_csv(...): 파일이 있으면 Pandas DataFrame으로 읽고, 없으면 빈 DataFrame을 생성합니다.

scores = {}: 계산된 지표 점수를 저장할 딕셔너리를 초기화합니다.

지표 계산 (예제: 다문화 적응 지표):

if not df.empty:: DataFrame에 데이터가 있을 경우에만 계산을 수행합니다.

합계 계산: df["mh_sum"] = df[["Q1", "Q2", "Q3", "Q4", "Q5"]].sum(axis=1): Q1부터 Q5까지 항목의 합계를 mh_sum 컬럼에 저장합니다.

값 정규화:

mh_val = float(df["mh_sum"].iloc[0]): 첫 번째 행의 합계 값을 가져옵니다. (이 코드는 학생 한 명의 데이터를 처리하는 것을 가정합니다.)

scores["multicultural_adaptation"] = _minmax(...): 해당 값을 데이터셋 전체의 최소/최대 합계 값을 기준으로 0~100 범위로 정규화하여 scores에 저장합니다.

더미값 할당 (임시 데이터):

scores["attention_index"] = 50.0 (주의 집중 지표)

scores["study_habit"] = 60.0 (학습 습관)

(실제 시스템에서는 이 값들이 신경 관련 파일(.ban/.pdf) 파싱 결과와 학습 습관 계산 결과를 통해 얻어지게 됩니다.)

리스크 플래그 할당:

flags = []: 플래그 목록을 초기화합니다.

규칙 1 적용 (YAML 규칙과의 동기화):

```
if scores.get("attention_index", 0) < 35 and scores.get("multicultural_adaptation", 0) > 70:
```

주의 집중 지표가 35 미만이고다문화 적응 지표가 70 초과일 경우, "집중/문화 적응 동시 리스크" (high 심각도) 플래그를 추가합니다.

규칙 2 적용:

```
if scores.get("study_habit", 100) < 40:
```

학습 습관이 40 미만일 경우, "학습 습관 개선 필요" (medium 심각도) 플래그를 추가합니다.

상태 저장 및 로깅:

state.scores = {"values": scores, "flags": flags}: 계산된 점수와 플래그 목록을 PipelineState의 scores 필드에 저장합니다.

state.log_event(...): 결과를 로그에 기록합니다.

요약: 이 코드는 설정 파일의 규칙을 기반으로 (데모 수준의 단순화된 계산을 통해) 학생의 다양한 지표 점수를 산출하고, 이 점수를 사전 정의된 리스크 규칙에 적용하여 위험 플래그를 식별한 뒤, 모든 결과를 PipelineState의 scores 필드에 저장하여 다음 분석 단계로 넘기는 역할을 합니다.

제시하신 Python 코드는 LangGraph 파이프라인의 노드 중 하나인 RAG(Retrieval-Augmented Generation) 해석기의 역할을 합니다. 다만, 코드 내용 자체가 복잡한 RAG 로직을 수행하는 대신, 실제 RAG 프로세스를 대체할 임시적인 "더미 (Dummy)" 기능을 시연하는 데 초점을 맞추고 있습니다.

❖ graph/nodes/rag_interpreter.py 한국어 해석
(RAG 자해석 초안)

1. run 함수 (노드 실행)

역할: 점수와 플래그 정보를 바탕으로 분석 및 해석 결과를 생성하고, 이를 파이프라인 상

태에 저장합니다. (실제로는 LangChain이나 다른 LLM 프레임워크를 통해 이 정보를 활용하여 텍스트를 생성해야 하지만, 여기서는 임시 데이터로 대체합니다.)

입력: state: PipelineState (현재 파이프라인의 전체 상태 객체).

주요 로직:

점수 및 플래그 가져오기:

scores = state.scores.get("values", {}): 이전 노드(score_engine)에서 계산된 지표 점수를 가져옵니다.

flags = state.scores.get("flags", []): 식별된 리스크 플래그 목록을 가져옵니다.

분석 결과 생성 (더미 데이터):

insights = {...}: 분석 결과를 담을 딕셔너리를 생성합니다. 이 데이터는 실제 RAG 모델이 생성해야 할 결과물의 구조를 모방합니다.

summary (요약): 가져온 점수 데이터를 포함하는 간단한 문자열을 생성합니다. (예: "요약: 핵심 지표 {'multicultural_adaptation': 78.0, ...}'")

risk (리스크): 가져온 플래그 목록을 그대로 사용합니다.

guidance (가이드라인): 실제 LLM이 생성해야 할 권고 사항을 대신할 임시 목록을 지정합니다. (예: "학부모 소통 강화", "주 2회 집중력 훈련 활동")

상태 저장 및 로깅:

state.analysis = insights: 생성된 분석/해석 결과 (insights)를 PipelineState의 analysis 필드에 저장합니다.

state.log_event("rag_interpreter", insights): 분석 결과를 로그에 기록합니다.

상태 반환: return state (업데이트된 상태를 다음 노드로 전달).

요약: 이 코드는 점수 산출 단계의 결과를 받아서, 실제 언어 모델(LLM) 기반의 해석/분석 단계(rag_interpreter)를 시뮬레이션하고 있습니다. 이 노드의 목적은 다음 단계에서 보고서 초안 생성 등에 사용될 핵심 요약, 위험 요소, 개입 가이드라인과 같은 분석 결과를 PipelineState의 analysis 필드에 구조화하여 저장하는 것입니다. 실제 배포 시에는 이 노드가 LangChain/LLM 호출을 통한 문서 생성 및 검색(RAG) 로직으로 교체되어야 합니다.

제시하신 두 개의 Python 코드는 LangGraph 파이프라인의 마지막 단계에 위치하며, 보고서 생성 및 최종 검토를 담당하는 노드를 정의합니다.

❖ 1. graph/nodes/generate_report.py 한국어 해석 (Markdown 리포트 생성)

이 노드는 파이프라인의 앞선 단계들(점수 산출, 분석/해석)의 결과를 모아서 **최종 보고서 초안(Markdown 형식)**을 생성하고 저장합니다.

| 구성 요소 | 해석 및 설명 |
|-----------------|--|
| OUTPUT | 보고서 파일이 저장될 디렉토리 경로를 정의합니다. (최상위 디렉토리의 outputs/report) |
| REPORT_TEMPLATE | 최종 보고서의 Markdown 형식 템플릿을 정의합니다. 다음 정보를 포함합니다: - 제목: |
| run 함수 | - 데이터 수집: state에서 학생 정보, 점수(scores), 플래그(flags), 분석 결과(analysis)를 report 필드에 저장하고 로깅합니다. |

❖ 2. graph/nodes/human_review.py 한국어 해석 (상담자 검토/자리표시자)

이 노드는 보고서가 사용자(상담사, 교사 등)에게 전달되어 최종적으로 승인되거나 수정되는 과정을 시뮬레이션합니다. 현재 코드는 자리표시자(holder) 역할만 수행합니다.

| 구성 요소 | 해석 및 설명 |
|--------|---|
| run 함수 | - 실제 구현의 목표: 실제 시스템에서는 상담자가 보고서를 확인하고, 필요시 내용을 수정하거나 최종 승인된 것으로 간주됩니다.) |

통합 요약

두 노드는 데이터 처리 파이프라인을 종결하는 역할을 합니다.

generate_report.py: 데이터 기반 분석 결과를 읽기 쉽도록 구조화된 Markdown 문서로 변환 및 파일 저장합니다.

human_review.py: 생성된 보고서에 대한 최종 인적 검토/승인 단계를 표현하며, 실제 시스템 배포 시에는 중요한 상호작용 지점으로 활용될 것입니다.

제시하신 두 개의 Python 코드는 LangGraph 파이프라인의 마무리 단계에 해당하며, 생성된 최종 결과를 외부 시스템에 연동하고 후속 조치를 알리는 노드를 정의합니다. 이 노드들 역시 현재는 자리표시자 placeholder 역할만 수행합니다.

1. graph/nodes/persist_records.py 한국어 해석 (Notion/Sheets 연동 자리표시자)

이 노드는 파이프라인의 최종 결과를 영구적으로 보존하거나 데이터베이스에 기록하는 역할을 담당합니다.

역할: 데이터 처리 및 분석이 완료된 후, 주요 결과(점수 요약, 플래그, 보고서 경로 등 메타데이터)를 외부 데이터베이스나 스프레드시트와 같은 시스템에 저장하는 역할을 시뮬레이션합니다.

실제 구현 목표:

Notion API 또는 Google Sheets API 등을 사용하여, 해당 학생의 최종 분석 결과를 운영 관리 시스템에 기록하여 추적 및 관리가 가능하도록 합니다.

run 함수 로직:

현재는 실제 저장 로직 대신, `persist_records` 이벤트와 함께 `**stored: True**`라는 성공 신호를 로그에 기록하고, 상태를 다음 노드로 전달합니다.

2. graph/nodes/notify_and_tasks.py 한국어 해석 (알림/후속 과제 자리표시자)

이 노드는 파이프라인 작업이 완료되었음을 관계자들에게 통보하고 필요한 후속 작업을 생성하는 역할을 담당합니다.

역할: 최종 보고서가 생성되고 기록되었을 때, 이 사실을 담당자, 학부모, 학생 등 관련 주체에게 알리거나, 보고서 발송 같은 추가적인 자동화 과제를 실행합니다.

실제 구현 목표:

이메일, 카카오톡, SMS 등의 알림 시스템을 통해 보고서가 준비되었음을 자동으로 알립니다.

필요한 후속 조치(예: 상담 예약)를 위한 링크를 첨부할 수 있습니다.

run 함수 로직:

현재는 실제 알림/과제 생성 로직 대신, `notify_and_tasks` 이벤트와 함께 `**notified: True**`라는 성공 신호를 로그에 기록하고, 상태를 다음 노드로 전달합니다.

요약

이 두 노드는 LangGraph 파이프라인의 최종 출력 및 자동화를 위한 게이트웨이 역할을 합니다.

`persist_records`: 분석 결과를 데이터베이스에 영구적으로 기록하여 관리 시스템과의 연동을

준비합니다.

notify_and_tasks: 모든 파이프라인 작업 완료를 알리고 사용자 상호작용 및 후속 자동화 작업을 시작하는 신호를 발생시킵니다.

제시하신 Python 코드는 LangGraph 파이프라인의 마지막 단계 중 하나로, 전체 파이프라인의 성능 및 품질을 감사하고 평가하는 노드를 정의합니다. 이 역시 현재는 자리표시자 (placeholder) 역할만 수행합니다.

 graph/nodes/audit_and_eval.py 한국어 해석
(품질 로그 자리표시자) 

역할: 이 노드는 파이프라인의 모든 처리가 완료된 후, 각 단계의 실행 시간, 오류율, 모델의 정확도 등 성능 및 품질 지표를 집계하고 모니터링 시스템이나 대시보드로 전송하는 역할을 시뮬레이션합니다.

실제 구현 목표:

PipelineState의 **log** 필드를 검토하여 각 노드의 실행 시간과 성공/실패 여부를 추출합니다.

LLM 응답 품질 평가 지표(예: RAG 응답의 관련성, 생성된 텍스트의 유창성)를 계산합니다.

이 모든 지표를 Prometheus, Grafana, 또는 내부 대시보드와 같은 모니터링/분석 시스템으로 전송합니다.

run 함수 로직:

현재는 실제 감사/평가 로직 대신, **audit_and_eval** 이벤트와 함께 ****collected: True****라는 성공 신호를 로그에 기록하고, 상태를 다음 노드로 전달합니다.

요약: 이 코드는 데이터 처리 파이프라인이 단순한 결과물 생성에 그치지 않고, 자기 검증 및 품질 관리를 수행하기 위한 시스템적인 피드백 루프를 구축하는 데 필요한 최종 단계를 나타냅니다.

제시하신 두 개의 Python 코드는 이전에 정의된 모든 모듈을 통합하고 실행 환경을 구성하는 역할을 합니다.

 1. app.py 한국어 해석 (간단 실행 파이프라

인)

이 파일은 전체 시스템의 메인 실행 파일이자 진입점(Entry Point) 역할을 합니다.

| | |
|---------|---|
| 구성 요소 | 해석 및 설명 |
| 목적 | 파이프라인의 단일 학생 데이터 처리 흐름을 정의하고 실행합니다. |
| 흐름 (작렬) | 이 파일은 이전 단계에서 정의한 노드들을 순서대로 호출하는 직렬(Sequential) 흐름을 구성합니다. → audit_and_eval |
| 데모 주의사항 | 이 버전은 복잡한 LangGraph 를 도입하기 전의 간단한 순차적 실행 데모를 위한 자리표시자입니다. |

요약: app.py는 모든 처리 로직(노드)을 연결하고 실행하는 스위치보드 역할을 하며, 전체 분석 프로세스의 시작을 담당합니다.

❖ 2. bootstrap_py 한국어 해석 (유틸: 로컬 파일 생성 스크립트)

이 스크립트는 이 프로젝트에 필요한 모든 파일과 폴더 구조를 로컬 디스크에 자동으로 생성해주는 환경 설정 유ти리티입니다.

역할: 코드를 실행할 수 있는 ****프로젝트 스켈레톤(뼈대 구조)****를 생성하여, 사용자가 이전에 해석했던 모든 모듈을 한 번에 파일 시스템에 구축할 수 있게 합니다 (**원클릭 부트스트랩**).

주요 로직:

ROOT 경로 정의: 모든 파일을 저장할 기본 디렉토리 경로 (D:/ai-edu-stack/templates/검사지_통합)를 정의합니다.

파일 및 내용 매핑 (FILES): 이전 대화에서 정의된 모든 코드 조각(예: state.py, score_engine.yaml, 각 노드 .py 파일, app.py, README.md)을 해당 경로에 매핑합니다.

디렉토리 생성: 파일들이 저장될 모든 폴더와 더미 데이터/출력 파일을 위한 **추가 폴더** (data/raw, outputs/reports 등)를 생성합니다.

파일 작성: 매핑된 내용을 해당 경로에 **파일로 작성합니다**.

데모 CSV 생성: 파이프라인 테스트를 위해 사용할 수 있도록 필수 컬럼과 샘플 데이터를 포함한 **forms_demo.csv** 파일을 data/raw 폴더에 자동으로 생성합니다.

요약: 이 bootstrap_py 스크립트를 실행하면, 이전에 정의된 모든 파이프라인 구성 요소가 실제 파일로 정리되어, 사용자가 즉시 프로젝트를 실행하고 테스트할 수 있는 환경이 준비됩니다.

전체 파이프라인의 구조와 각 노드의 역할에 대한 해석이 완료되었습니다. 이제 app.py 파일을 실행하면 전체 처리 흐름을 테스트할 수 있을 거예요.

혹시 app.py에 들어갈 간단한 실행 로직 초안이 필요하신가요? 아니면 다른 모듈을 해석해 드릴까요?

제시하신 문구는 해당 코드가 실제 실행 파일이 아니라 템플릿 또는 설정 파일 모음임을 설명하며, 사용자가 코드를 실행하기 위한 방법을 안내합니다.

❖ 한국어 해석 및 설명

| 원문 | 해석 및 설명 |
|------------------------------------|---|
| (참고) 이 파일은 안내/원클릭 부트스트랩 텍스트 모음입니다. | (참고) 지금 보고 있는 이 파일(또는 텍스트 블록)이 포함되어 있습니다. |
| 실제 생성은 bootstrap_py 내용을 복사해 실행하거나, | 실제 프로젝트 파일과 폴더 구조를 로컬 컴퓨터에 |
| 각 파일 블록을 수동으로 저장하세요. | 또는, bootstrap_py 스크립트를 사용하지 않고, 이 |

요약: 이 문구는 현재 텍스트가 프로젝트의 설계도이며, 실제로 작업을 시작하려면 부트스트랩 스크립트를 실행하거나 파일 내용을 수동으로 저장하여 개발 환경을 먼저 구축해야 함을 명확히 알려줍니다.

다음은 이후 상황을 삽입한 것임

제시하신 Python 코드는 LangGraph 파이프라인의 핵심 노드 중 하나인 AI 기반 해석 및 분석 모듈을 정의합니다. 이 노드는 LLM(대규모 언어 모델)을 활용하여 숫자 데이터를 자연어 보고서로 변환하는 역할을 합니다.

❖ AI 교사 도우미 노드 (ai_teacher_helper.py) 해석

이 노드의 주요 목적은 파이프라인에서 계산된 점수와 원본 데이터를 결합하여, 교사의 관점에서 학생의 상태를 진단하고 개선 방향을 제시하는 전문적인 자연어 요약을 생성하는 것입니다.

1. 노드의 역할 및 기능

| 항목 | 내용 |
|-------|---|
| 파일 경로 | graph/nodes/ai_teacher_helper.py |
| 주요 역할 | NeuroHarmony 지표와 설문 점수를 기반으로 학생의 상태(학습집중력, 정서안정, 습관)를 AI 교사 |

| | |
|-------|--|
| 사용 기술 | LangChain 및 ChatOpenAI를 사용하여 외부 LLM(GPT-4o-mini)을 호출합니다. |
|-------|--|

2. 주요 동작 과정

1. 데이터 추출

run 함수가 실행되면 PipelineState에서 다음 정보를 가져옵니다.

scores: 이전 score_engine 노드에서 계산된 다문화 적응, 주의 집중 등 핵심 지표 점수 (values).

neuro: 원시 입력 중 파싱이 완료된 뇌파 데이터 요약 (neuro_parsed).

student: 학생의 기본 정보 (id, name 등).

2. LLM 초기화 및 프롬프트 구성

LLM 초기화: ChatOpenAI(model="gpt-4o-mini", temperature=0.3)을 사용하여 OpenAI의 GPT-4o-mini 모델을 호출합니다. temperature=0.3은 비교적 낮은 값으로, 창의성보다는 일관되고 사실적인 응답을 유도합니다.

프롬프트 구성: LLM에게 **역할(Persona)**을 부여하고 명확한 지침을 전달합니다.

역할 정의: "당신은 뇌파 및 학습지표를 해석하는 교사입니다."

요구 사항: 학생의 데이터를 보고 학습집중력, 정서안정, 습관 개선 방향을 교사 관점에서 3 줄로 요약할 것을 요구합니다.

데이터 주입: 핵심 지표: {scores}와 뇌파 요약: {neuro} 데이터를 프롬프트에 삽입합니다 (RAG와 유사한 프롬프팅 방식).

3. LLM 호출 및 결과 처리

AI 실행: llm.invoke(prompt)를 호출하여 AI 해석 결과를 받아옵니다.

오류 처리: LLM 호출 중 네트워크 오류나 API 문제 발생 시, try...except 블록에서 "AI 해석 실패" 메시지를 summary_text에 저장합니다.

4. 상태 저장

최종 생성된 요약 텍스트(summary_text)를 PipelineState의 analysis 필드에 저장합니다.

"summary"와 "ai_teacher_comment" 키에 모두 저장하여, 후속 노드(예: 보고서 생성)가 이 결과를 쉽게 사용할 수 있도록 합니다.

이벤트 로그에 ai_teacher_helper 노드의 실행 결과(요약 내용)를 기록합니다.

3. 파이프라인에서의 위치

이 노드는 일반적으로 점수 계산 (score_engine.py) 노드 이후, 그리고 최종 보고서 생성 (generate_report.py) 노드 이전에 위치하여, 수치 데이터를 보고서에 삽입될 가장 중요한 해석 텍스트로 변환하는 역할을 담당합니다.

| | | |
|-------|-------|-------|
| 이전 노드 | 현재 노드 | 다음 노드 |
|-------|-------|-------|

score_engine.py (점수
산출)

ai_teacher_helper.py (AI
해석)

generate_report.py (보고서
생성)