

- SBI(Startup Brain Index) 진단 시스템
 - 시스템 개요
 - 4대 영역
 - 설치 방법
 - 1. 의존성 설치
 - 2. 파일 구조 확인
 - 사용 방법
 - 1. API 서버 실행
 - 2. API 엔드포인트
 - GET /
 - GET /health
 - GET /items
 - POST /analyze-sbi (Step 2: 설문 + 가상 뇌파 통합 지수)
 - POST /analyze-combined (기존 뇌파 결합)
 - POST /submit-survey
 - 3. 테스트 실행
 - 주요 기능
 - Step 2: 뇌파 데이터 결합 분석 엔진
 - 1. 데이터 구조화
 - 2. 문항 필터링
 - 3. 채점 로직
 - 4. 영역별 점수 계산
 - Step 3-1: 지식 DB 구축 및 스크래핑 (무료 로컬)
 - 파일 구조
 - 테스트 결과
 - API 문서
 - 주의사항
 - 개발자 정보

SBI(Startup Brain Index) 진단 시스템

뇌교육학 박사 이중환의 'SBI(Startup Brain Index)' 진단 시스템입니다.

시스템 개요

이 시스템은 96개의 설문 문항을 기반으로 창업가의 뇌지수를 진단합니다.

4대 영역

1. 창업공감 및 동기부여
2. 창업위기감수 및 극복
3. 창업두뇌활용 및 계발
4. 주체적책임 및 창업의식

설치 방법

1. 의존성 설치

```
pip install -r requirements.txt
```

2. 파일 구조 확인

다음 파일이 필요합니다:

- **survey_items.csv**: 설문 문항 데이터 (96개 문항)

사용 방법

1. API 서버 실행

```
python main.py
```

또는

```
uvicorn main:app --reload
```

서버가 실행되면 <http://localhost:8000>에서 API를 사용할 수 있습니다.

2. API 엔드포인트

GET /

시스템 정보 및 사용 가능한 엔드포인트 목록

GET /health

시스템 상태 확인 및 데이터 검증

GET /items

전체 문항 정보 조회

POST /analyze-sbi (Step 2: 설문 + 가상 뇌파 통합 지수)

설문 응답만 보내면 가상 뇌파(MockEEGProvider) 를 자동 매칭해 영역별 가중치로 SBI 통합 지수를 산출합니다. 설문-뇌파 차이 20% 이상이면 `inconsistency_flag: true`로 리포트 권장.

요청 예시:

```
{
  "responses": { "1": 5, "2": 4, ... },
  "excluded_sequences": []
}
```

응답 필드: `survey_전체평균`, `survey_정규화_0_100`, `eeg_영역별` (motivation, resilience, innovation, responsibility), `영역별_통합점수`, `통합지수_0_100`, `inconsistency_flag`, `message` 등

POST /analyze-combined (기존 뇌파 결합)

설문(SBI) 응답과 뇌파(EEG) 지표를 결합하여 통합 지수(0~100)를 산출합니다. `brainwave` 를 생략하면 설문만으로 종합 지수를 반환합니다.

요청 예시:

```
{
  "responses": { "1": 5, "2": 4, ... },
  "excluded_sequences": [],
  "brainwave": { "engagement": 80, "focus": 70 },
  "survey_weight": 0.6,
  "eeg_weight": 0.4
}
```

응답 필드: `survey_sbi_score`, `survey_sbi_normalized`, `eeg_engagement`, `eeg_focus`, `combined_index`, 영역별 점수, `message` 등

POST /submit-survey

설문 응답 제출 및 점수 계산

요청 예시:

```
{
  "responses": {
    "1": 5,
    "2": 4,
    "3": 5,
    "4": 3,
    "5": 5,
    ...
  },
  "excluded_sequences": [1, 2, 3]
}
```

응답 예시:

```
{
  "전체평균": 3.9,
  "영역별점수": [
    {
      "영역명": "창업공감 및 동기부여",
      "평균점수": 4.33,
      "문항수": 24,
      "포함된_순번": [1, 2, 3, ...]
    },
    ...
  ],
  "사용된_문항수": 96,
  "제외된_순번": [],
  "메시지": "총 96개 문항으로 점수 계산 완료"
}
```

3. 테스트 실행

```
python test_scoring.py
```

주요 기능

Step 2: 뇌파 데이터 결합 분석 엔진

- 설문(SBI) 점수와 뇌파(EEG) 지표(engagement, focus, alpha, beta 등)를 결합
- 0~100 통합 지수(**combined_index**) 산출, 가중치(survey_weight, eeg_weight) 지원
- 뇌파 미제공 시 설문만으로 정규화된 종합 지수 반환

1. 데이터 구조화

- CSV 파일의 '영역', '하위역량', '하위요소', '문항(보정)' 컬럼 매핑
- 4대 영역 클래스 정의

2. 문항 필터링

- **get_filtered_score** 함수를 통해 특정 순번 문항 제외 가능
- 예: 96문항 중 24개 제외하여 점수 계산

3. 채점 로직

- 모든 응답은 1~5점 리커트 척도
- 역채점 문항 없음
- 단순 평균 계산

4. 영역별 점수 계산

- 각 영역별 평균 점수 계산
- 영역별 사용된 문항 수 추적

Step 3-1: 지식 DB 구축 및 스크래핑 (무료 로컬)

- **블로그:** `blog_scraper.py` — requests + BeautifulSoup으로 장산뇌혁신데이터랩 블로그 최신 포스트 제목·본문 수집
- **유튜브 자막:** `youtube_transcript.py` — youtube-transcript-api로 영상 ID 입력 시 자막 텍스트 추출
- **로컬 DB:** `knowledge_db.py` — SQLite `sbi_knowledge.db`, 테이블 `id`, `source_type`(블로그/유튜브), `title`, `content`, `url`, `created_at`, url 기준 중복 방지
- **검색:** TfidfVectorizer(또는 키워드 매칭)로 역량 키워드 검색 시 상위 3건 반환
- **검증:** `python test_knowledge_step3.py` — 블로그 1건·유튜브 1건 샘플 수집 후 DB 저장 및 키워드 검색 확인

자세한 검증 결과는 `STEP3_1_TEST_REPORT.md` 참고.

파일 구조

```
project2_StartupBrainIndex/
├── main.py                # FastAPI 메인 애플리케이션
├── models.py              # 데이터 모델 (SBI + 뇌파)
├── data_loader.py         # CSV 데이터 로더
├── scoring.py             # 채점 로직
├── analysis_engine.py     # Step 2: 뇌파 결합 분석 엔진
├── eeg_provider.py        # Step 2: Mock 뇌파 제공
├── blog_scraper.py        # Step 3-1: 블로그 수집
├── youtube_transcript.py  # Step 3-1: 유튜브 자막 수집
├── knowledge_db.py        # Step 3-1: 로컬 지식 DB 및 검색
├── test_scoring.py        # 설문/채점 유닛 테스트
├── test_combined_sbi.py   # Step 2 통합 지수 테스트
├── test_knowledge_step3.py # Step 3-1 지식 DB 검증
├── requirements.txt       # 의존성 목록
├── survey_items.csv       # 설문 문항 데이터
├── sbi_knowledge.db       # 지식 DB (실행 후 생성)
└── README.md             # 이 파일
```

테스트 결과

모든 테스트가 성공적으로 통과했습니다. 자세한 내용은 `test_report.md`를 참조하세요.

API 문서

서버 실행 후 다음 URL에서 자동 생성된 API 문서를 확인할 수 있습니다:

- Swagger UI: <http://localhost:8000/docs>
- ReDoc: <http://localhost:8000/redoc>

주의사항

- CSV 파일은 **cp949** 인코딩으로 저장되어야 합니다.
- 응답 점수는 반드시 1~5 사이의 정수여야 합니다.
- 제외할 순번은 1~96 사이의 정수여야 합니다.

개발자 정보

이 시스템은 뇌교육학 박사 이중환의 SBI 진단 시스템을 기반으로 구축되었습니다.