

# 암웨이 뉴트리라이트 제품 추천 시스템 데이터베이스 명세서

## 1. 개요

### 1.1 목적

LLM+RAG 기반 챗봇을 통한 개인 맞춤형 건강기능식품 추천 시스템 구축을 위한 데이터베이스 설계

### 1.2 시스템 구성

- 데이터베이스: PostgreSQL 14+
- 확장 모듈: pgvector (벡터 검색), JSONB (JSON 데이터 처리)
- 인코딩: UTF-8
- 타임존: UTC

### 1.3 핵심 기능

- 사용자 건강 데이터 기반 제품 추천
- 과학적 근거 기반 추천 설명
- 벡터 검색을 통한 유사 제품 탐색
- 추천 성능 모니터링 및 개선

## 2. 테이블 상세 명세

### 2.1 제품 관련 테이블

#### 2.1.1 products (제품 기본 정보)

컬럼명	데이터타입	제약조건	설명
product_id	SERIAL	PRIMARY KEY	제품 고유 ID
product_name	VARCHAR(100)	NOT NULL	제품명
product_code	VARCHAR(50)	UNIQUE	제품 코드
category	VARCHAR(50)	NOT NULL	제품 카테고리 (건강기능식품, 일반식품 등)
subcategory	VARCHAR(50)		세부 카테고리 (기초영양, 영양건강, 기능성)
food_type	VARCHAR(100)		식품유형
manufacturer	VARCHAR(100)		제조사
is_global	BOOLEAN	DEFAULT FALSE	글로벌/로컬 제품 구분
is_active	BOOLEAN	DEFAULT TRUE	활성 상태
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시
updated_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	수정일시

인덱스: idx\_products\_category (category, subcategory)

2.1.2 product\_details (제품 상세 정보)

컬럼명	데이터타입	제약조건	설명
id	SERIAL	PRIMARY KEY	고유 ID
product_id	INTEGER	FOREIGN KEY	제품 ID (products 참조)
intake_method	TEXT		섭취 방법
precautions	TEXT		주의사항
nutritional_info	JSONB		영양성분 정보
ingredients_list	TEXT		원재료명
calories_per_serving	DECIMAL(6,2)		1회 제공량당 칼로리
serving_size	VARCHAR(50)		1회 제공량
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

2.1.3 allergen\_info (알레르기 정보)

컬럼명	데이터타입	제약조건	설명
id	SERIAL	PRIMARY KEY	고유 ID
product_id	INTEGER	FOREIGN KEY	제품 ID (products 참조)
allergen_name	VARCHAR(50)	NOT NULL	알레르기 유발 물질명
allergen_type	VARCHAR(20)		알레르기 유형 (함유, 혼입가능성)
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

인덱스: idx\_allergen\_info\_product (product\_id)

2.2 성분 및 기능성 테이블

2.2.1 functional\_ingredients (기능성 원료/성분)

컬럼명	데이터타입	제약조건	설명
ingredient_id	SERIAL	PRIMARY KEY	성분 고유 ID
ingredient_name_kr	VARCHAR(100)	NOT NULL	성분명 (한국어)
ingredient_name_en	VARCHAR(100)		성분명 (영어)
ingredient_type	VARCHAR(50)		성분 유형 (비타민, 미네랄, 식물추출물 등)
description	TEXT		성분 설명
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

2.2.2 product\_ingredients (제품-원료 관계)

컬럼명	데이터타입	제약조건	설명
id	SERIAL	PRIMARY KEY	고유 ID
product_id	INTEGER	FOREIGN KEY	제품 ID (products 참조)
ingredient_id	INTEGER	FOREIGN KEY	성분 ID (functional_ingredients 참조)
amount	VARCHAR(50)		함량 정보
unit	VARCHAR(20)		단위 (mg, g, IU 등)
daily_value_percentage	DECIMAL(5,2)		일일권장량 대비 백분율
is_main_ingredient	BOOLEAN	DEFAULT FALSE	주요 성분 여부
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

인덱스: idx\_product\_ingredients\_product (product\_id)

## 2.3 건강 영역 관련 테이블

### 2.3.1 health\_areas (건강 영역 마스터)

컬럼명	데이터타입	제약조건	설명
area_id	SERIAL	PRIMARY KEY	건강영역 고유 ID
area_name	VARCHAR(100)	NOT NULL UNIQUE	건강영역명
area_category	VARCHAR(50)		영역 카테고리
description	TEXT		영역 설명
is_kfda_approved	BOOLEAN	DEFAULT TRUE	식약처 인정 기능성 여부
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

### 2.3.2 product\_health\_functions (제품-건강영역 관계)

컬럼명	데이터타입	제약조건	설명
id	SERIAL	PRIMARY KEY	고유 ID
product_id	INTEGER	FOREIGN KEY	제품 ID (products 참조)
area_id	INTEGER	FOREIGN KEY	건강영역 ID (health_areas 참조)
function_description	TEXT		구체적인 기능성 설명
evidence_level	VARCHAR(20)		근거 수준
is_primary_function	BOOLEAN	DEFAULT FALSE	주요 기능성 여부
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

인덱스:

- idx\_product\_health\_functions\_product (product\_id)
- idx\_product\_health\_functions\_area (area\_id)

2.4 건강 지표 및 점수 테이블

2.4.1 health\_indicators (건강 지표 마스터)

컬럼명	데이터타입	제약조건	설명
indicator_id	SERIAL	PRIMARY KEY	건강지표 고유 ID
indicator_name	VARCHAR(100)	NOT NULL	건강지표명
indicator_category	VARCHAR(50)		지표 카테고리 (체성분, 혈액검사 등)
measurement_unit	VARCHAR(20)		측정 단위
normal_range_min	DECIMAL(10,2)		정상 범위 최솟값
normal_range_max	DECIMAL(10,2)		정상 범위 최댓값
description	TEXT		지표 설명
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

2.4.2 health\_scores (건강 점수 마스터)

컬럼명	데이터타입	제약조건	설명
score_id	SERIAL	PRIMARY KEY	건강점수 고유 ID
score_name	VARCHAR(100)	NOT NULL	건강점수명
score_category	VARCHAR(50)		점수 카테고리
description	TEXT		점수 설명
calculation_method	TEXT		계산 방법
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

2.4.3 indicator\_health\_impact (건강지표-영향영역 관계)

컬럼명	데이터타입	제약조건	설명
id	SERIAL	PRIMARY KEY	고유 ID
indicator_id	INTEGER	FOREIGN KEY	건강지표 ID (health_indicators 참조)
area_id	INTEGER	FOREIGN KEY	건강영역 ID (health_areas 참조)
impact_level	VARCHAR(20)	NOT NULL	영향 수준 (주의, 관리)
correlation_strength	DECIMAL(3,2)		연관성 강도 (0.0-1.0)
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

인덱스:

- idx\_indicator\_health\_impact\_indicator (indicator\_id)
- idx\_indicator\_health\_impact\_area (area\_id)

2.4.4 score\_health\_impact (건강점수-영향영역 관계)

컬럼명	데이터타입	제약조건	설명
id	SERIAL	PRIMARY KEY	고유 ID
score_id	INTEGER	FOREIGN KEY	건강점수 ID (health_scores 참조)
area_id	INTEGER	FOREIGN KEY	건강영역 ID (health_areas 참조)
impact_level	VARCHAR(20)	NOT NULL	영향 수준 (주의, 관리)
priority_weight	DECIMAL(3,2)		우선순위 가중치
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

인덱스:

- idx\_score\_health\_impact\_score (score\_id)
- idx\_score\_health\_impact\_area (area\_id)

2.5 추천 시스템 관련 테이블

2.5.1 functional\_combinations (기능성 조합 분류)

컬럼명	데이터타입	제약조건	설명
combination_id	SERIAL	PRIMARY KEY	조합 고유 ID
combination_name	VARCHAR(100)	NOT NULL	조합명
antioxidant_function	BOOLEAN	DEFAULT FALSE	항산화 기능 포함 여부
metabolic_function	BOOLEAN	DEFAULT FALSE	대사 기능 포함 여부
muscle_function	BOOLEAN	DEFAULT FALSE	근육 기능 포함 여부
description	TEXT		조합 설명
target_health_scores	JSONB		대상 건강점수 목록
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

2.5.2 product\_functional\_combinations (제품-기능성조합 관계)

컬럼명	데이터타입	제약조건	설명
id	SERIAL	PRIMARY KEY	고유 ID
product_id	INTEGER	FOREIGN KEY	제품 ID (products 참조)
combination_id	INTEGER	FOREIGN KEY	조합 ID (functional_combinations 참조)
primary_mechanism	TEXT		주요 작용 기전
supporting_evidence	TEXT		지원 근거
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

2.5.3 recommendation\_rules (제품 추천 규칙)

컬럼명	데이터타입	제약조건	설명
rule_id	SERIAL	PRIMARY KEY	규칙 고유 ID
rule_name	VARCHAR(100)	NOT NULL	규칙명
condition_type	VARCHAR(50)		조건 유형 (건강지표기반, 건강점수기반 등)
condition_criteria	JSONB		조건 (JSON 형태)
recommended_products	JSONB		추천 제품 목록
rule_priority	INTEGER	DEFAULT 1	규칙 우선순위
rule_description	TEXT		규칙 설명
is_active	BOOLEAN	DEFAULT TRUE	활성 상태
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

2.5.4 lifestyle\_nutrition\_mapping (생활습관-영양소 매핑)

컬럼명	데이터타입	제약조건	설명
id	SERIAL	PRIMARY KEY	고유 ID
lifestyle_category	VARCHAR(50)		생활습관 카테고리 (식생활, 운동, 수면)
lifestyle_item	VARCHAR(100)		생활습관 항목
required_nutrients	JSONB		필요 영양소 목록
recommended_products	JSONB		추천 제품
nutrition_rationale	TEXT		추천 근거
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

2.6 과학적 근거 테이블

2.6.1 scientific\_evidence (과학적 근거)

컬럼명	데이터타입	제약조건	설명
evidence_id	SERIAL	PRIMARY KEY	근거 고유 ID
ingredient_id	INTEGER	FOREIGN KEY	성분 ID (functional_ingredients 참조)
area_id	INTEGER	FOREIGN KEY	건강영역 ID (health_areas 참조)
study_title	TEXT		연구 제목
journal_name	VARCHAR(200)		저널명
publication_year	INTEGER		발행 년도
doi	VARCHAR(100)		DOI
pmid	VARCHAR(20)		PubMed ID
study_type	VARCHAR(50)		연구 유형 (RCT, 메타분석 등)
evidence_level	VARCHAR(20)		근거 수준 (A, B, C 등급)
key_findings	TEXT		주요 발견사항
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

2.7 RAG 시스템 지원 테이블

2.7.1 product\_embeddings (제품 벡터 임베딩)

컬럼명	데이터타입	제약조건	설명
id	SERIAL	PRIMARY KEY	고유 ID
product_id	INTEGER	FOREIGN KEY	제품 ID (products 참조)
content_type	VARCHAR(50)		콘텐츠 유형 (제품설명, 기능성, 성분 등)
content_text	TEXT	NOT NULL	텍스트 내용
embedding	vector(1536)		벡터 임베딩 (OpenAI embedding)
metadata	JSONB		추가 메타데이터
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

인덱스:

- idx\_product\_embeddings\_product (product\_id)
- idx\_product\_embeddings\_vector (embedding vector\_cosine\_ops) - pgvector 확장 필요

2.8 로그 및 모니터링 테이블

2.8.1 recommendation\_logs (추천 로그)

컬럼명	데이터타입	제약조건	설명
log_id	SERIAL	PRIMARY KEY	로그 고유 ID
session_id	VARCHAR(100)		세션 ID
user_health_data	JSONB		사용자 건강 데이터
recommended_products	JSONB		추천된 제품들
recommendation_reason	TEXT		추천 이유
user_feedback	INTEGER		사용자 피드백 점수
llm_model_version	VARCHAR(50)		LLM 모델 버전
created_at	TIMESTAMP	DEFAULT CURRENT_TIMESTAMP	생성일시

인덱스: idx\_recommendation\_logs\_session (session\_id)

### 3. 관계도

#### 3.1 핵심 관계

products (1) ↔ (N) product\_ingredients (N) ↔ (1) functional\_ingredients  
products (1) ↔ (N) product\_health\_functions (N) ↔ (1) health\_areas  
health\_indicators (1) ↔ (N) indicator\_health\_impact (N) ↔ (1) health\_areas  
health\_scores (1) ↔ (N) score\_health\_impact (N) ↔ (1) health\_areas  
products (1) ↔ (N) product\_functional\_combinations (N) ↔ (1) functional\_combinations

#### 3.2 데이터 흐름

- 1. 사용자 건강 데이터 입력 → health\_indicators/health\_scores 매핑
- 2. 영향 영역 도출 → indicator\_health\_impact/score\_health\_impact 조회
- 3. 제품 후보 필터링 → product\_health\_functions 조회
- 4. 기능성 조합 분석 → product\_functional\_combinations 조회
- 5. 최종 추천 생성 → recommendation\_rules 적용
- 6. 근거 자료 제시 → scientific\_evidence 조회

### 4. JSON 스키마 정의

#### 4.1 nutritional\_info (영양성분 정보)

json



```
{
  "calories": 120,
  "protein": 15.5,
  "fat": 2.3,
  "carbohydrate": 8.1,
  "sodium": 45,
  "vitamins": {
    "vitamin_c": {"amount": 500, "unit": "mg", "daily_value_percent": 556}
  },
  "minerals": {
    "calcium": {"amount": 200, "unit": "mg", "daily_value_percent": 25}
  }
}
```

## 4.2 condition\_criteria (추천 조건)

```
json

{
  "health_scores": {
    "노화억제분석지수": {"status": "주의", "threshold": 60}
  },
  "health_indicators": {
    "체지방률": {"status": "관리", "value_range": [25, 30]}
  },
  "lifestyle": {
    "exercise_frequency": "low",
    "dietary_pattern": "high_carb"
  }
}
```

## 4.3 user\_health\_data (사용자 건강 데이터)

```
json
```

```
{
  "health_scores": {
    "노화억제분석지수": 58,
    "만성질환억제분석지수": 72,
    "근육밸런스지수": 65
  },
  "health_indicators": {
    "체중": 75.5,
    "허리둘레": 88,
    "공복혈당": 105,
    "체지방률": 23.2
  },
  "lifestyle": {
    "exercise_frequency": "medium",
    "sleep_hours": 6.5,
    "stress_level": "high"
  }
}
```

## 5. 인덱스 전략

### 5.1 기본 인덱스

- 모든 기본키에 자동 인덱스 생성
- 외래키에 인덱스 생성으로 조인 성능 최적화

### 5.2 복합 인덱스

- products 테이블: (category, subcategory)
- 조회 패턴에 따른 다중 컬럼 인덱스

### 5.3 벡터 인덱스

- pgvector 확장을 통한 유사도 검색 최적화
- IVFFlat 알고리즘 사용

## 6. 성능 최적화 고려사항

### 6.1 파티셔닝

- recommendation\_logs 테이블을 날짜 기준으로 파티션 분할 고려
- 대용량 데이터 처리 시 적용

### 6.2 JSONB 최적화

- GIN 인덱스를 통한 JSON 필드 검색 성능 향상

- 자주 조회되는 JSON 경로에 표현식 인덱스 생성

## 6.3 캐싱 전략

- 자주 조회되는 제품-건강영역 매핑 정보 캐싱
- Redis를 통한 추천 결과 임시 저장

## 7. 보안 및 프라이버시

### 7.1 데이터 보호

- 개인 건강 정보는 세션 단위로만 저장
- 민감 정보 암호화 처리
- 정기적인 로그 데이터 익명화

### 7.2 접근 제어

- 역할 기반 접근 제어 (RBAC) 구현
- API 레벨에서 데이터 접근 제한

### 7.3 감사 추적

- 모든 데이터 변경 이력 추적
- 시스템 접근 로그 관리

## 8. 확장성 고려사항

### 8.1 수평 확장

- 읽기 전용 복제본을 통한 조회 성능 향상
- 제품 데이터와 로그 데이터 분리

### 8.2 마이크로서비스 대응

- 도메인별 데이터베이스 분리 가능한 구조
- API 기반 데이터 접근

### 8.3 버전 관리

- 스키마 변경 이력 관리
- 점진적 마이그레이션 지원