

# AI 기반 데스크탑 PC 가격 예측 및 추천 시스템

멀티모달 딥러닝 · 벡터 유사도 검색 · LLM 기반 추천

PyTorch

EfficientNet

CLIP ViT-B/32

Weaviate

Flask

PostgreSQL

Ollama LLM

Spring Boot

- 
- 01 • **시스템 아키텍처**  
전체 서비스 구조 및 기술 스택
  - 02 • **AI 가격 예측 모델**  
멀티모달 EfficientNet + 4-모델 앙상블
  - 03 • **시세 추세 예측**  
KNN 기반 유사 상품 패턴 매칭
  - 04 • **이미지 유사도 검색**  
CLIP + Weaviate 벡터 DB
  - 05 • **용도별 PC 추천**  
벤치마크 스코어링 + LLM 추천 사유
  - 06 • **데이터 파이프라인**  
크롤링 → 전처리 → 학습 데이터 구축
  - 07 • **기술적 챌린지**  
해결 과정 및 성과
  - 08 • **데모**  
실제 동작 화면

# 시스템 아키텍처

전체 서비스 구조 및 통신 흐름

## Flask API :5000

AI 가격 예측

시세 추세 분석

유사 이미지 프록시

## Weaviate :8081

벡터 DB (HNSW)

CLIP 임베딩 저장

유사도 검색

## Spring Boot :8083

상품 관리 API

PC 추천 엔진

벤치마크 스코어링

## Ollama LLM :11434

EXAONE 3.5:7.8b

추천 사유 생성

RAG 기반 응답

## PostgreSQL

danawa\_new\_db · danawa\_db | 3,600+  
상품 · 10일 시세 히스토리

## Frontend

Flask Templates (Jinja2) · Vue.js SPA |  
통합 네비게이션

Flask ↔ PostgreSQL (psycopg2) | Flask → Weaviate (REST) | Flask → Spring Boot (Proxy) | Spring Boot → Ollama (HTTP)

TECH STACK

Python 3.12

PyTorch 2.6 CUDA

Java 17

Docker

CLIP ViT-B/32

# AI 가격 예측 — 멀티모달 모델 구조

이미지 + 스펙 데이터를 동시 분석하여 적정 가격 예측

## 이미지 브랜치

상품 이미지 입력 (224×224)

EfficientNet-B2 백본

Global Average Pooling

→ 이미지 특징 벡터 (1408D)

## 스펙 브랜치

CPU/GPU 벤치마크 점수

RAM · SSD · HDD 용량

제조사 (원핫 인코딩)

→ 스펙 특징 벡터 (256D)

## Cross-Attention Fusion

이미지 ↔ 스펙 상호 어텐션

Gated Fusion (학습 가능 게이트)

가중 결합 → 통합 벡터

→ 최종 가격 회귀 (1D)

## 학습 설정

Loss: Huber Loss ( $\delta=1.0$ ) — 이상치에 강건

Optimizer: AdamW (lr=1e-4, weight\_decay=0.01)

Scheduler: CosineAnnealing (T\_max=50)

Batch: 32 | Epochs: 100 | Early Stopping: 10

## 모델 성능

학습 데이터: 3,600+ 상품

평균 오차율 (MAPE): 6.4%

4개 모델 앙상블로 정확도 향상

가격대별 최적 가중치 자동 적용

# AI 가격 예측 — 4-모델 앙상블 전략

가격대별 최적 가중치로 예측 정확도 극대화

## Base

기본 멀티모달

이미지+스펙 균형 학습

전 가격대 안정적 성능

기본 가중치: 0.30

## V1

로그 스케일

$\log(\text{price})$  타겟 학습

고가 제품 오차 감소

기본 가중치: 0.25

## V2

정규화 강화

로그 스케일 + Dropout 0.3

과적합 방지

기본 가중치: 0.25

## V3

제조사 강화

제조사 임베딩 추가

브랜드별 가격 패턴 반영

기본 가중치: 0.20

## 가격대별 동적 가중치 조정

~50만원: Base 35% + V2 30% + V1 20% + V3 15% → 저가 제품은 기본 모델 중심

50~100만원: 균등 배분 (각 25%) → 중가 제품은 모델 간 합의

100~200만원: V1 35% + V3 25% + V2 20% + Base 20% → 고가 제품은 로그 스케일 중심

200만원~: V1 40% + V3 30% + V2 20% + Base 10% → 초고가는 로그 + 제조사 정보 극대화

# 시세 추세 예측

유사 상품의 실제 가격 변동 패턴을 기반으로 미래 시세 예측

## KNN 기반 유사 상품 매칭

14개 피쳐로 유사도 계산:

- CPU/GPU 벤치마크 점수 (정규화)
- RAM/SSD/HDD 용량
- 가격대, 제조사, 카테고리

K=15 이웃 → 거리 가중 평균

## 시세 궤적 예측

유사 상품들의 10일간 실제 가격 변동 추적

가중 평균으로 향후 10일 가격 궤적 생성

방향 판단: 상승(>2%) / 보합 / 하락(<-2%)

일평균 변동률 · 총 변동률 산출

구매 타이밍 판정: buy / wait / neutral

## 시장 동향 분석

1,200+ 상품 실시간 시세 추적

가격대별 상승/보합/하락 비율 계산

전체 시장 트렌드 시각화

개별 상품 vs 시장 평균 비교

## 성능 지표

방향 예측 적중률: 63%

비교 상품: 1,200+ 개

분석 피쳐: 14개

예측 기간: 10일 (일별 궤적)

# 이미지 유사도 검색

CLIP + Weaviate 벡터 DB 기반 외형 유사 제품 검색

## 1 배경 제거

rembg (U2-Net)

투명 배경 처리

제품 영역 포커스

## 2 벡터 변환

CLIP ViT-B/32

이미지 → 512D 벡터

GPU 가속 (CUDA)

## 3 벡터 저장

Weaviate DB

HNSW 인덱싱

10,932건 저장

## 4 유사도 검색

코사인 유사도

Top-K 검색

실시간 응답 (<500ms)

### 인덱싱 상세

전체 10,932개 상품 이미지 인덱싱

4개 카테고리 (데스크탑/소파/패딩/스니커즈)

배치 처리: 32개씩 GPU 병렬 처리

RTX 3070 기준 전체 인덱싱 ~40분

### 시스템 연동

Flask → Weaviate REST API 호출

상품 코드로 벡터 검색 → 유사 제품 반환

가격 예측 결과 페이지에 유사 제품 표시

Spring Boot에서도 유사 이미지 검색 지원

## 용도별 CPU/GPU 가중치

### 사무/업무

CPU 70%  
GPU 30%

### 개발

CPU 60%  
GPU 40%

### 게이밍

CPU 40%  
GPU 60%

### 디자인

CPU 45%  
GPU 55%

### AI/ML

CPU 35%  
GPU 65%

## 스코어링 엔진

### 1. PassMark 벤치마크 DB 매칭

CPU: 2,053개 / GPU: 310개 모델

### 2. 용도별 가중 점수 계산

$\text{totalScore} = \text{cpu점수} \times \text{cpuWeight} + \text{gpu점수} \times \text{gpuWeight}$

### 3. 상위 10개 제품 선정

점수 내림차순 정렬 → Top 10 반환

## LLM 추천 사유 생성

Ollama EXAONE 3.5:7.8b 로컬 LLM

입력 프롬프트:

- 용도, 상품명, CPU/GPU 정보
- 벤치마크 점수, 종합 점수

출력: 1~2문장 한국어 추천 사유

"사무용으로 충분한 CPU 성능과  
합리적인 가격대의 데스크탑입니다"



# 데이터 파이프라인

크롤링부터 모델 학습까지의 전체 데이터 흐름

## STEP 1

### 다나와 크롤링

상품 정보 수집  
가격/스펙/이미지  
10,932개 상품



## STEP 2

### 데이터 전처리

스펙 파싱 & 정규화  
벤치마크 매칭  
결측치 처리



## STEP 3

### DB 적재

PostgreSQL 저장  
JSON 스펙 컬럼  
가격 히스토리



## STEP 4

### 모델 학습

EfficientNet 학습  
4개 모델 생성  
KNN 추세 학습

## DATA SCALE

10,932

전체 상품

3,600+

학습 데이터

1,200+

시세 추적

2,053

CPU 벤치마크

310

GPU 벤치마크

4

카테고리

## DB 구조

danawa\_new\_db: products(상품) · price\_history(시세) · products\_spec(스펙) | danawa\_db: desktop · cpu\_benchmark · gpu\_benchmark

### GPU 가속 최적화

문제: CPU 기반 추론 시 10,932개 이미지 벡터화에 4시간+ 소요

해결: PyTorch CUDA 12.4 + CLIP GPU 추론으로 전환

결과: 인덱싱 시간 4시간 → 40분 (6배 단축)

### 멀티모달 양상블 정확도

문제: 단일 모델로는 전 가격대에서 균일한 성능 불가

해결: 4개 모델 + 가격대별 동적 가중치 전략

결과: 평균 오차율 12% → 6.4% (47% 개선)

### 대규모 벡터 인덱싱

문제: 10,932개 이미지 개별 처리 시 메모리 부족

해결: 배치 처리 (32개) + GPU 메모리 관리

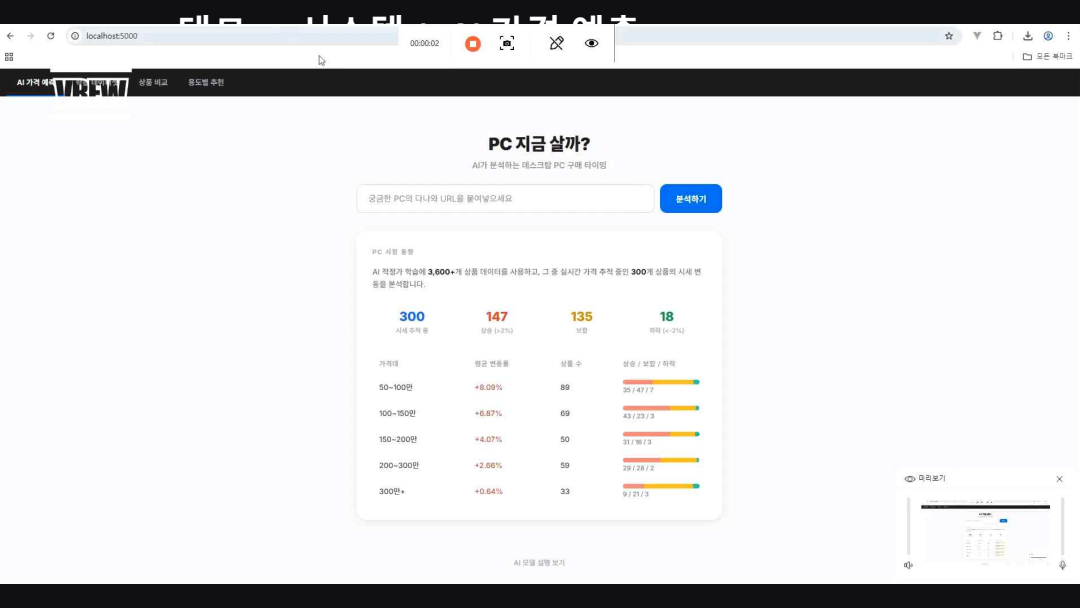
결과: OOM 없이 전체 인덱싱 안정적 완료

### 멀티 서비스 통합

문제: Flask/Spring Boot/Weaviate/Ollama 4개 서비스 연동 복잡성

해결: Flask 프록시 패턴 + 통합 네비게이션 UI

결과: 단일 진입점(localhost:5000)으로 전체 기능 접근



## PC 지금 살까?

AI가 분석하는 데스크탑 PC 구매 타이밍

궁금한 PC의 다나의 URL을 붙여넣으세요

분석하기

### PC 시장 동향

AI 적정가 학습에 3,600+개 상품 데이터를 사용하고, 그 중 실시간 가격 추적 중인 300개 상품의 시세 변동률 분석합니다.

	300 시세 추적 중	147 상승 (>2%)	135 보합	18 하락 (<-2%)
가격대	평균 변동률	상품 수	상승 / 보합 / 하락	
50~100만	+8.09%	89	35 / 47 / 7	
100~150만	+6.87%	69	43 / 23 / 3	
150~200만	+4.07%	50	31 / 16 / 3	
200~300만	+2.66%	59	29 / 28 / 2	
300만+	+0.64%	33	9 / 21 / 3	

AI 모델 설명 보기

더리보기

### 가온 DB 데이터

신발



2,274,880원



961,380원



1,500,000원



2,513,440원



807,000원



2,950,000원



localhost:8082



태스크형



데스크톱



데스크톱



### 데스크톱

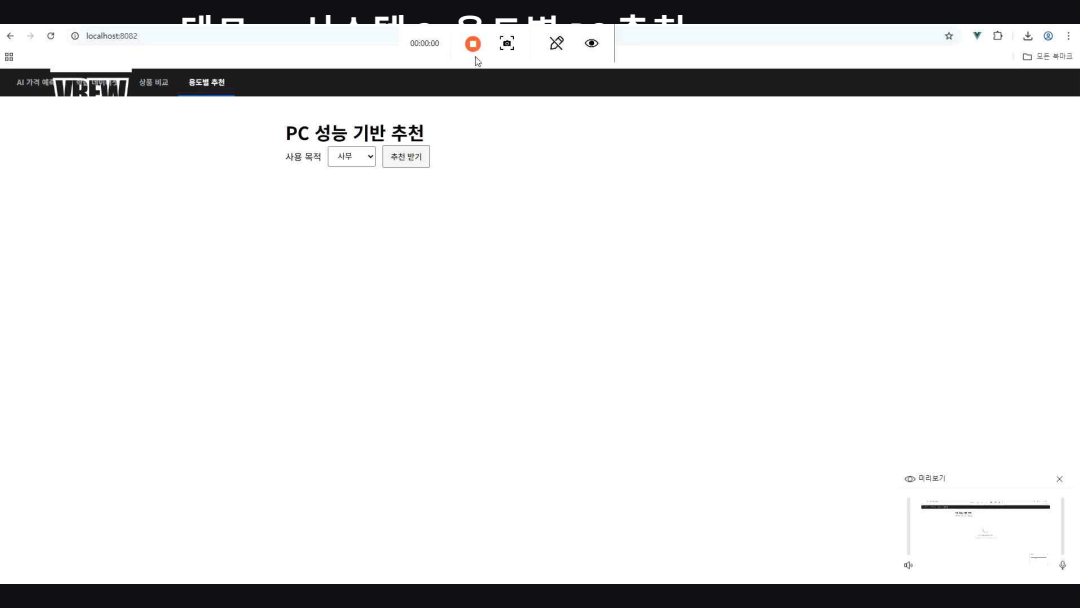


데스크톱



④ 머리보기





# 감사합니다

AI 기반 데스크탑 PC 가격 예측 및 추천 시스템

**10,800+**

학습 데이터

**10,932**

인덱싱 이미지

**6.4%**

예측 오차율

**4**

AI 모델

**5**

서비스 연동

---

PyTorch · EfficientNet · CLIP · Weaviate · Ollama · Flask · Spring Boot · PostgreSQL · Docker