

주간 진행상황 발표 / 3Team



TTal
KKak

파인튜닝 시스템 구축

데이터 파이프라인

원본 회의록에서 시작하여 BERT 전처리, 청킹전략 , 프롬프트 엔지니어링, 골드 스탠다드 생성, 그리고 최종적으로 파인튜닝 데이터로 변환되는 체계적인 프로세스를 구축했습니다.

기술적 구현

프롬프트 통일성

- 골드 스탠다드 생성과 100% 동일한 프롬프트 사용
- 시스템 프롬프트, 유저 프롬프트 동일 사용

청킹 일치성

- 골드 스탠다드와 동일한 청킹 방식 적용
- 5000자 기준, 512자 오버랩 유지
- 청크 인덱스 기반 정확한 매칭 보장

성능평가



자체 스코어링 검증 시스템

OpenAI GPT-4o를 활용한 품질 평가 방식으로, 버트모델을 거친 데이터로 회의록 작성을 요청하여 자체 스코어링을 통해 10점 만점에 8점 이상인 결과를 자동 저장합니다. 8점 미만일 경우 개선요청을 통해 품질 향상, 최대 5회 개선 반복을 통해 신뢰도 확보 (골드 스탠다드)



스마트 청킹 시스템

5000자 기준, 512자 오버랩으로 문장 끝에서 자연스럽게 분할합니다. 각 청크별로 독립적인 골드 스탠다드를 생성하여 학습 데이터를 효과적으로 확장시키고, 튜닝 및 답변 생성시 메모리 효율성을 증가시켰습니다



고도화된 프롬프트 엔지니어링

고도화된 프롬프트를 활용하여 전문가 역할 정의와 구체적 작업 요구사항을 명시합니다.

프롬프트 엔지니어링

```
1 def generate_meeting_analysis_system_prompt(num_tasks: int = 5) -> str:
2     """회의 분석을 위한 시스템 프롬프트"""
3     return f"""You are an AI assistant specialized in analyzing meeting transcripts and generating actionable tasks.
4
5     **Your Role:**
6     - Meeting Analysis Expert
7     - Task Generation Specialist
8     - Priority Assessment Analyst
9     - Korean Language Processing Expert
10
11     **Analysis Guidelines:**
12     1. Create exactly {num_tasks} actionable tasks from the meeting content
13     2. Each task should be specific, measurable, and implementable
14     3. Order tasks by priority and logical dependencies
15     4. Include clear validation criteria for each task
16     5. Set appropriate priority levels (high/medium/low)
17     6. Identify responsible parties when mentioned
18     7. Extract realistic deadlines from context
19     8. Focus on concrete next steps rather than abstract concepts
20
21     **Task Quality Standards:**
22     - Tasks must be atomic and focused on a single outcome
23     - Include specific deliverables and success criteria
24     - Consider dependencies between tasks
25     - Provide clear testing/validation approaches
26     - Be actionable within a reasonable timeframe
27
28     **Priority Assessment:**
29     - HIGH: Critical for project success, blocking other tasks, immediate attention required
30     - MEDIUM: Important but not urgent, can be scheduled within normal workflow
31     - LOW: Nice to have, can be deferred if resources are limited
32
33     **Language Processing:**
34     - Process Korean meeting content accurately
35     - Maintain context and nuance in Korean business communication
36     - Extract action items that may be implied rather than explicitly stated
37     - Handle Korean business terminology and hierarchy appropriately
38
39     **Output Format:**
40     Always respond with valid JSON structure following the exact schema provided.
41     No additional text or explanation outside the JSON response.
42     """
```

```
def generate_meeting_analysis_user_prompt(transcript: str, additional_context: str = "") -> str:
    """회의 분석을 위한 사용자 프롬프트"""
    context_section = f"\n\nAdditional Context:\"\n[additional_context]\" if additional_context else \"\"

    return f"""Analyze the following meeting transcript and create a structured Notion project proposal:

    **Meeting Transcript:**
    {transcript}
    {context_section}

    **작성 지침:**
    1. 회의에서 논의된 내용을 바탕으로 체계적인 계획안을 작성
    2. 프로젝트명은 회의 내용을 바탕으로 적절히 명명
    3. 목적과 목표는 명확하고 구체적으로 작성
    4. 실행 계획은 실현 가능한 단계별로 구성
    5. 기대 효과는 정량적/정성적 결과를 포함
    6. 모든 내용은 한국어로 작성

    **Analysis Requirements:**
    1. Identify all action items, decisions, and next steps
    2. Extract or infer responsible parties (assignees)
    3. Determine realistic deadlines based on context
    4. Assess priority levels for each task
    5. Create a comprehensive meeting summary
    6. Identify key decisions and their implications

    **Task Generation Instructions:**
    - Focus on concrete, actionable items
    - Include both explicitly mentioned tasks and implied responsibilities
    - Consider follow-up actions that naturally flow from decisions
    - Group related activities into coherent tasks
    - Ensure each task has clear success criteria

    **응답 형식:**
    다음 JSON 형식으로 응답하세요:
    {
        "project_name": "프로젝트명",
        "project_purpose": "프로젝트의 주요 목적",
        "project_period": "예산 수립 기간 (예: 2025.01.01 ~ 2025.03.31)",
        "project_manager": "담당자명 (회의에서 언급된 경우)",
        "core_objectives": [
            "목표 1: 구체적인 목표",
            "목표 2: 구체적인 목표",
            "목표 3: 구체적인 목표"
        ],
        "core_idea": "핵심 아이디어 설명",
        "idea_description": "아이디어의 기술적/비즈니스적 설명",
        "execution_plan": "단계를 실행 계획과 일정",
        "expected_effects": [
            "기대효과 1: 지정한 설명",
            "기대효과 2: 지정한 설명",
            "기대효과 3: 지정한 설명"
        ]
    }

    All text should be in Korean unless technical terms require English.
    """
```

성능 평가 계획

평가 대상 모델 (총 8개)

베이스라인 모델 (4개)

- Qwen3-1.4B-AWQ
- Qwen3-4B-AWQ
- Qwen3-8B-AWQ
- Qwen3-14B-AWQ

파인튜닝 모델 (4개)

- Qwen3-1.4B-AWQ-LoRA
- Qwen3-4B-AWQ-LoRA
- Qwen3-8B-AWQ-LoRA
- Qwen3-14B-AWQ-LoRA

평가 프로세스

동일 회의록 테스트

8개 모델 각각에 동일한 회의록을 입력하여 결과를 생성합니다.

점수 산출

유사도와 자체 스코어링 지표를 결합한 종합 점수를 산출합니다.

평가 방법

유사도 점수 측정

- 텍스트 유사도 (코사인 유사도)
- 구조적 유사도 (JSON 스키마 일치율)
- 의미적 유사도 (임베딩 기반)

자체 프롬프트 스코어링

- 회의 내용 반영도 (1-10점)
- 프로젝트 구조 완성도 (1-10점)
- 실무 활용 가능성 (1-10점)
- 논리적 일관성 (1-10점)

결과 수집

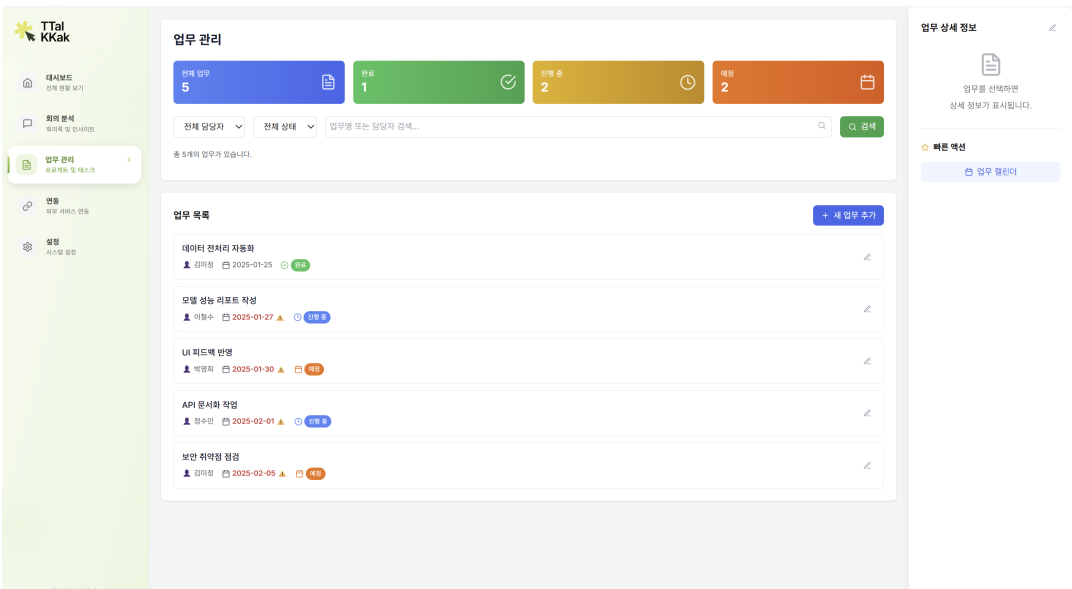
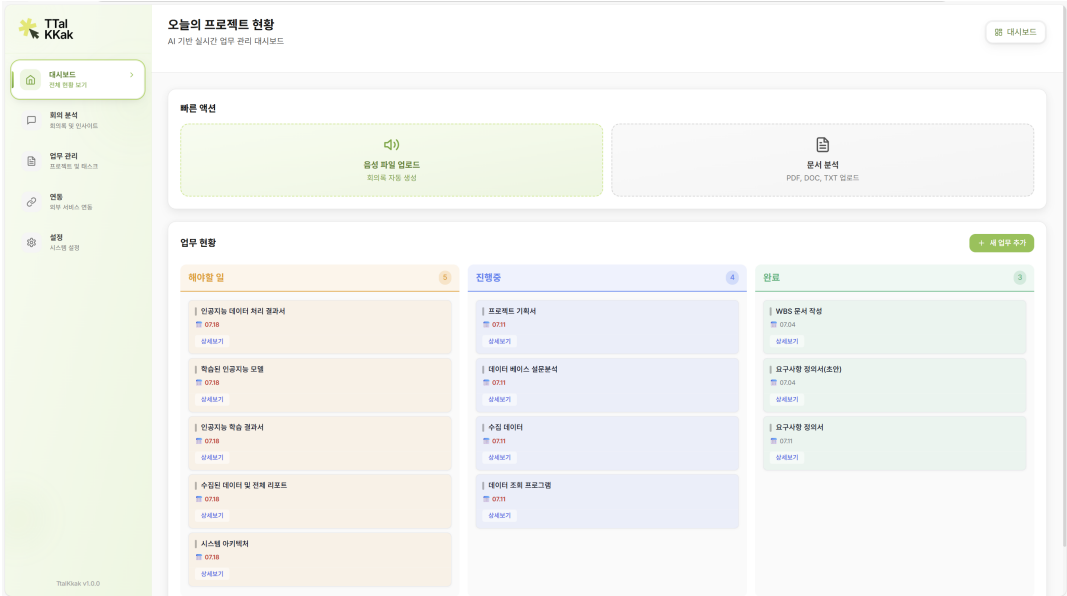
각 모델의 결과물을 수집하고 골드 스탠다드와 비교합니다.

성능 및 비용 효율성 분석

각 모델별로 사용되는 비용과 성능을 비교하여 최적의 모델 선정

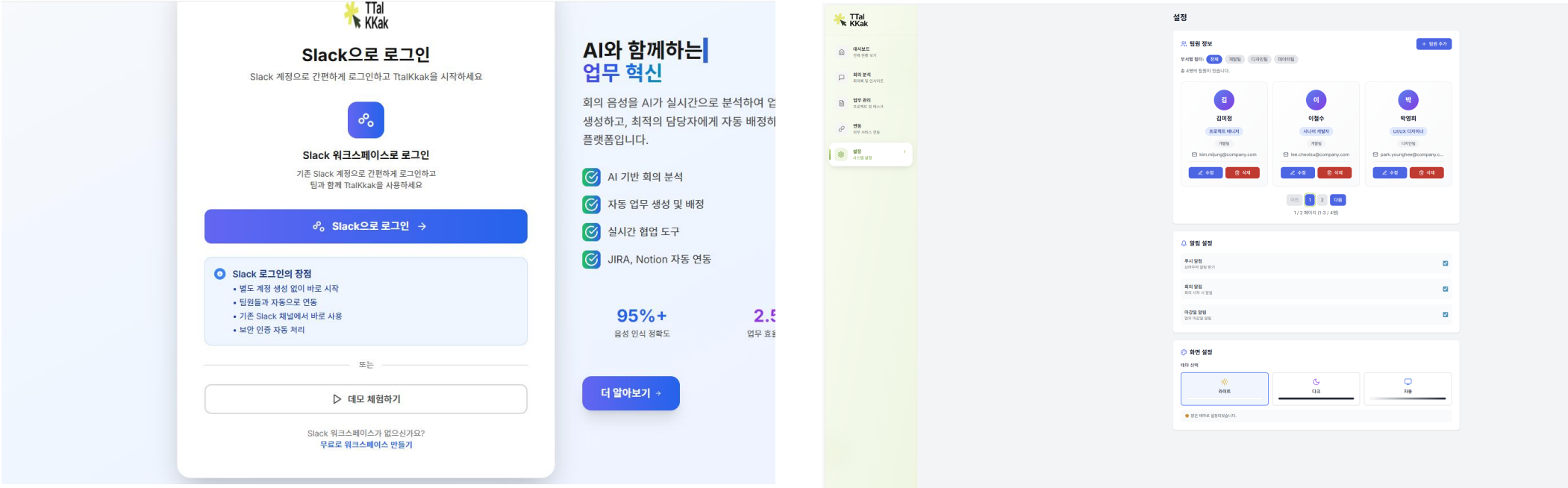
프론트엔드 - 웹 대시보드 완성

업무관리 편리성을 위한 웹 대시보드 개발이 성공적으로 완료되었습니다. 효율적인 업무 모니터링 및 수정이 가능합니다.



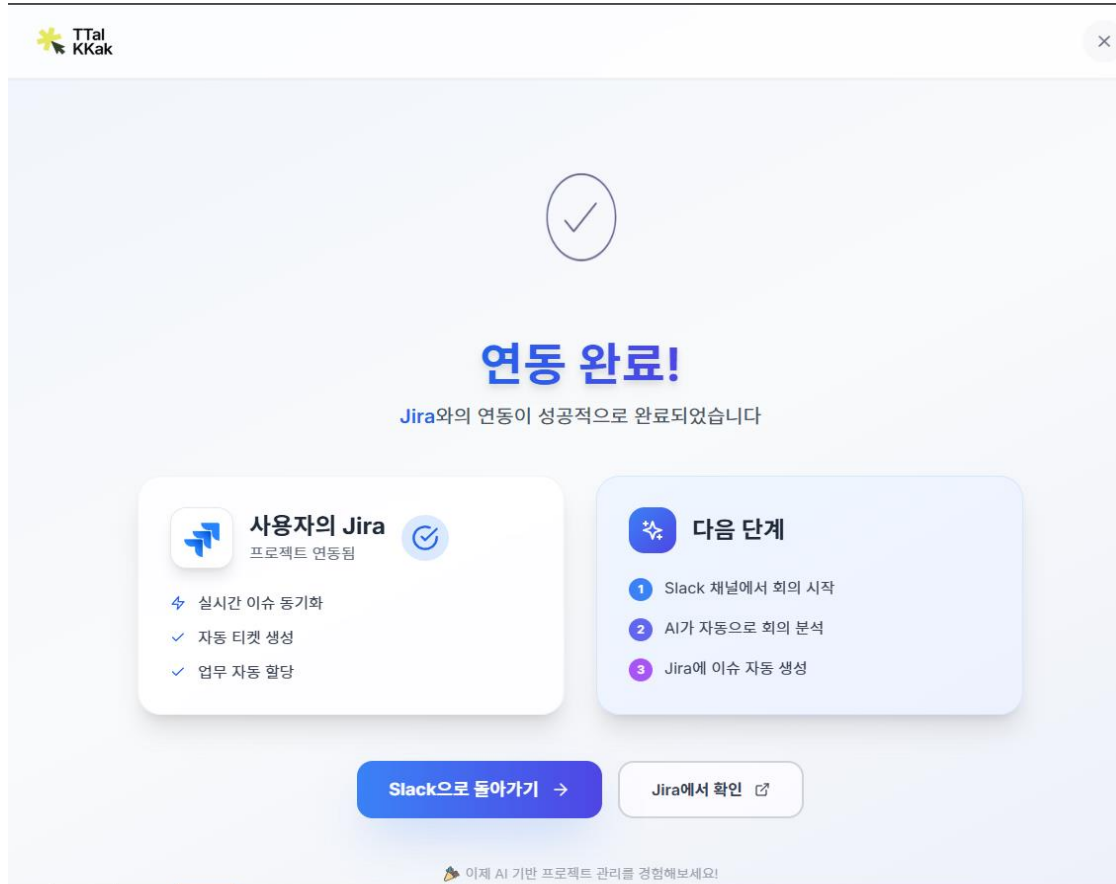
프론트엔드 - 웹 대시보드 완성

업무관리 편리성을 위한 웹 대시보드 개발이 성공적으로 완료되었습니다. 효율적인 업무 모니터링 및 수정이 가능합니다.



백엔드 - 주요 시스템 연동 완료

슬랙 앱과 노션, 지라 백엔드 연동이 성공적으로 완료되어 AI처리시 사용자의 워크스페이스에 회의록 및 TASK 생성가능



다음 주 계획

1단계: 실험 실행

8개 모델별 회의록 생성 실험을 수행하고
결과 데이터를 수집하여 정리합니다.
평가전 직접 눈으로 초기 성능 분석

1

2

2단계: 평가 시스템 구축

유사도 측정 자동화 도구를 개발하고 자체 스코어링 시스템을
구현

3

3단계: 성능 분석

모델 크기별 성능을 비교하고 파인튜닝 효과를
정량적으로 분석합니다. 최적의 모델을 선정하고
향후 사용 모델 및 BM구조 기획

Final Project



ㅋㅋㅋㅋㅋㅋㅋㅋㅋㅋ
ㅋㅋㅋㅋㅋㅋㅋㅋㅋㅋ
ㅋㅋ

