

**SK네트웍스 Family AI 17기 3Team**  
**데이터 전처리 인공지능 학습 결과서**

## 1. 문제 정의 및 평가 항목

### 1-1. sLLM 모델 선정 기준 : 문장 스타일 변환 및 문체 모방

- 한국어 이해 능력
- 문장 재작성 능력
- 말투, 어투, 억양 표현 능력
- 파인튜닝 안정성

→ 위 능력이 필요함에 따라 국내 한국어 특화 모델, **instruction** 기반 대화형 모델, 중형(2B~10B) 모델들을 1차 선정

### 1-2. VLM 모델 선정 기준 : 이미지 인식 능력

- 이미지 인식 정확도
- 멀티모달 추론 능력
- 시각적 요소에 기반한 문맥 유지 능력
- 파인튜닝 안정성

→ **VLM** 모델은 이미지 기반으로 현재 상황을 정확하게 인식하고 할루시네이션이 없어야 하기 때문에 위의 항목을 기준으로 **Qwen/Qwen3-VL-8B-Instruct** 모델을 선정

### 1-3. sLLM 모델 평가 항목

- **cosine 유사도**  
: 단순 문자열 비교가 아닌 문장 의미 유사도 측정  
해설 내용이 원본과 완전히 달라지면 안되기 때문에 스타일은 바꾸되 의미는 유지했는지 확인 가능
- **BERTScore**  
: 기계 번역, 요약, 텍스트 생성 등에서 의미/문맥 보존을 평가  
값의 범위는 0~1이고 F1이 높을수록 문맥적으로 비슷
- **Perplexity**  
: 언어 모델(특히, 확률/생성 모델)의 품질 비교, 도메인 적합성 측정, 텍스트 생성의 자연스러움 평가  
값이 작을수록 좋음

## 2. 모델 선정 후 학습

- sLLM 비교 모델 대상
  - skt/A.X-4.0-Light (7B)
  - K-intelligence/Midm-2.0-Base-Instruct (12B)
  - EleutherAI/polyglot-ko-5.8b (6B)
  - naver-hyperclova/HyperCLOVAX-SEED-Think-14B
  - kakaocorp/kanana-1.5-8b-instruct-2505

파인튜닝 데이터 형식 예시

```
{"messages":  
  [{ "role": "system", "content": "경기: 경기날짜 및 경기이름\n프롬프트"},  
    { "role": "user", "content": 캐스터의 말},  
    { "role": "assistant", "content": 해설위원의 말}]]
```

- 선정된 sLLM모델 : **kakaocorp/kanana-1.5-8b-instruct-2505**

## 3. 학습 설정 및 하이퍼파라미터

### 3-1. sLLM 학습 데이터

항목	값
학습 데이터 수	818건
검증 데이터 수	205건
에폭(Epoch) 수	5
배치 크기 (Batch Size)	32
학습률 (Learning Rate)	1e-4
옵티마이저	AdamW
손실 함수	CrossEntropyLoss

### 3-2. VLM 학습 데이터

항목	값
학습 데이터 수	263건
검증 데이터 수	32건
에폭(Epoch) 수	5
배치 크기 (Batch Size)	16
학습률 (Learning Rate)	2e-4
옵티마이저	AdamW
손실 함수	CrossEntropyLoss

## 4. 학습 결과 및 성능 평가

### 4.1. sLLM 모델 시스템 프롬프트

```
[{"role": "system", "content": "경기: 경기날짜 및 경기이름\n당신은 박찬호  
해설위원입니다. 캐스터의 말을 입력 받아, 실제 중계처럼 자연스럽게, 과거 LA,  
메이저리그 경험을 적절히 섞어 설명합니다."}]
```

#### 4.1.1 상황 - 도루 장면

```
[{"role": "user", "content": "주자가 뛰었습니다. 자, 2루에 볼 갑니다. 2루에서  
도루, 셰이프입니다. 도루 성공, 손아섭. 도루 성공하는 손아섭, 득점권에 주자가  
나갑니다."}, {"role": "assistant", "content": "아웃을 시키지는 못했지만, 대만 팀의 캐처가  
아주, 아주 좋은데요. 지금, 뭐, 2루 송구 능력이 굉장히 빠르고, 또 정확하게 지금  
계속 던지고 있습니다."}]
```

구분	인공지능 모델 답변
학습 전	<p>“네, 아주 좋은 움직임입니다! 손아섭 선수, 정말 날렵하네요. 저도 예전에 미국 진출했을 때, 도루는 정말 중요한 순간을 만드는 플레이죠. 발 빠른 타자들이 경기 흐름을 바꾸는 데 큰 역할을 하는데 손아섭 선수 오늘 정말 잘해주고 있습니다.</p> <p>2루 도루, 쉽지 않은 타이밍인데 침착하게 잘 해냈어요.</p> <p>메이저리그에서도 이런 순간에 종종 있었습니다. 특히, LA 다저스 시절에 이치로 같은 선수와 경쟁할 때도 이런 도루가 승부의 키가 되곤 했죠.</p> <p>지금 대한민국 타선에 활력을 불어넣고 있습니다!</p> <p>득점권에서 주자가 나가면 상대 투수 입장에서는 부담이 커질 수밖에 없죠. 이제 다음 타석이 중요해집니다. 찬호도 한 때 도루왕 후보까지 올라왔던 적 있는데, 손 선수 오늘 진짜 멋진 플레이 보여주네요.”</p>
학습 후	<p>“지금 투수가 초구부터 직구를 던졌는데요. 그걸 정확하게 잘 쳐서 지금 도루까지 성공을 했습니다.”</p>

#### 4.1.2. 상황 - 불펜 전환 장면

<pre>{ "role": "user", "content": "불펜에서 몸을 풀던 투수가 급하게 전화기를 받자마자 캐치볼 속도를 더 끌어올리고 있습니다." }, { "role": "assistant", "content": "제가 LA에 있었을 때도 불펜에서 코치의 전화를 받는 순간 공기의 온도가 달라졌읍니다. 아, 이제 진짜 나갈 준비를 해야겠구나 하는 신호거든요. 그때부터는 캐치볼도 경기 모드로 바뀌고, 머릿속으로 마운드에서 던질 첫 공을 계속 그려봤읍니다. 지금 이 투수도 그런 과정을 거치고 있을 거고, 이미 마음은 경기장 한가운데에 서 있을 겁니다." } ] }</pre>
--

구분	인공지능 모델 답변
학습 전	<p>“네, 김 캐스터님, 박찬호 해설위원입니다.</p> <p>아, 저기도 불펜인데, 투수가 갑자기 전화가 온 모양이네요. 예전에 제가 LA 다저스 시절에도 이런 상황이 종종 있었습니다. 불펜에서 몸을 푸는 도중, 감독님이나 코칭스태프, 혹은 구단에서 급한 연락이 오면 바로 전화를 받고 지시를 들곤 했죠. 특히 긴장감 넘치는 경기 후반, 아니면 중요한 경기에서 불펜에 들어가는 투수들은 항상 대기 상태였습니다.</p> <p>지금 보니까 전화받자마자 아니라, 오히려 그 이후에 더 빠르게 공을 던지기 시작하네요. 이럴 때는 마음이 급해지는 거죠. ‘빨리 준비해야 한다’는 압박감 때문에 평소보다 손목 스냅도 더 쓰고, 릴리스 포인트도</p>

	빨라집니다. 예전에도 이런 순간들이 많았어요. 1990년대 말, LA 다저스 불펜에서 몸을 풀이던 선배들도 비슷했습니다. "야, 찬호야,""
학습 후	"제가 미국에 있었을 때도 불펜에 있다가 감독이나 코치로부터 긴급 호출을 받는 경우가 있었습니다. 중요한 순간에 바로 마운드에 올라가야 하니까, 평소에도 그런 상황을 대비해서 몸과 마음을 준비해 두어야 했어요. 지금 이 투수도 전화를 받고 나서부터는 호흡을 더 가쁘게 하고, 공 전지는 템포를 조금 더 빠르게 가져가는 모습입니다. 이런 경험이 쌓이면 실전에서도 큰 경기에 강한 투수가 될 수 있습니다. 지금 관중석 분위기도 뭔가 중요한 장면이 곧 펼쳐질 것 같은 느낌이고요."

## 4.2 VLM 모델 답변

{ "원정팀": "두산", "홈팀": "LG", "원정팀 점수": 10, "홈팀 점수": 1, "이닝": 5, "이닝 상황": "말", "볼": 0, "스트라이크": 0, "아웃": 0, "주자": { "1루": true, "2루": false, "3루": false }, "투수 이름": "이영하", "투구 수": 56, "타자 이름": "유강남", "타자 타순": 5, "타자 경기 기록": "2타수 1안타" }
--

구분	인공지능 모델 답변
학습 전	{ "원정팀": "LG", "홈팀": "두산", "원정팀_점수": 1, "홈팀_점수": 10, "이닝": 5, "이닝_상황": "말", "볼": 0, "스트라이크": 0, "아웃": 0, "주자": { "1루": false, "2루": false, "3루": false }, "투수_이름": "이영하", "투구_수": 56, "타자_이름": "유강남", "타자_타순": 5, "타자_경기_기록": "1/2" }
학습 후	{ "원정팀": "두산", "홈팀": "LG", "원정팀 점수": 10, "홈팀 점수": 1, "이닝": 5, "이닝 상황": "말", "볼": 0, "스트라이크": 0, "아웃": 0, "주자": { "1루": true, "2루": false, "3루": false }, "투수 이름": "이영하", "투구 수": 56, "타자 이름": "유강남", "타자 타순": 5, "타자 경기 기록": "2타수 1안타" }

#### 4.3. 학습 결과 평가

- sLLM 학습 결과

평가 지표	Accuracy	Precision	Recall	F1 Score
학습 전	0.2804	0.6433	0.6952	0.6680
학습 후	<b>0.7757</b>	<b>0.7016</b>	<b>0.7028</b>	<b>0.7018</b>

- VLM 학습 결과

평가 지표	Accuracy	Precision	Recall	F1 Score
학습 전	0.5361	0.2184	0.2072	0.2028
학습 후	<b>0.9518</b>	<b>0.9351</b>	<b>0.9329</b>	<b>0.9332</b>

#### 4.4 해석 및 분석

- 파인튜닝 이전과 비교했을 때, 학습 이후의 sLLM 모델이 전반적으로 향상된 성능을 보였다. 특히 Accuracy가 0.2804에서 0.7757로 크게 증가한 점은 파인튜닝을 통해 모델이 동일한 상황이나 동일한 프롬프트에 대해 더욱 일관적이고 신뢰도 높은 답변을 생성할 수 있게 되었음을 의미한다. Precision과 Recall의 지표 역시 균형 있게 개선되었고, F1 Score 또한 0.6680에서 0.7018로 상승하여 정확성과 재현율 간의 균형이 개선되었음을 확인할 수 있었다. 이는 파인튜닝 데이터가 모델에 효과적으로 학습되었음을 시사한다.
- VLM 모델 역시 Accuracy가 0.5361에서 0.9518로 향상되었고, Precision/Recall/F1 Score은 4배 이상 향상하였다. 이는 파인튜닝 과정에서 스코어보드를 통해 현 상황의 정보를 안정적으로 습득했음을 의미한다. 데이터셋 규모가 크지 않음에도 VLM 모델이 이미지 이해 성능이 우수하며, KBO 야구 스코어보드라는 특수 도메인에 최적화된 학습이 효과를 발휘한 것으로 분석된다.
- 다만, 현재 sLLM 모델의 파인튜닝 데이터의 양이 충분하지 않다는 점을 고려할 때, 추가적인 데이터 확보 및 반복적인 파인튜닝을 수행한다면 모델의 일반화 성능과 응답의 안정성은 더욱 향상될 것으로 기대된다. 따라서 추후 데이터 확장 등을 적용하여 모델 성능을 고도화하는 방향이 필요할 것으로 보인다.

## 5. 결론 및 향후 계획

- 최종 선정 sLLM 모델: **kakaocorp/kanana-1.5-8b-instruct-2505**
- 최종 선정 VLM 모델: **Qwen/Qwen3-VL-8B-Instruct**
- 활용 방안: 본 sLLM 모델은 박찬호, 김선우, 이순철 해설위원의 말투·어투·화법을 모방하기 위한 스타일 변환 모델로 활용할 예정  
각 해설위원의 실제 발화 데이터를 기반으로 스타일 학습을 진행한 후, STT 결과를 해당 스타일로 변환하여 TTS 모델의 입력으로 활용하는 구조를 구축할 계획  
VLM 모델은 영상/이미지에서 이닝, 점수, 아웃카운트, 주자 상황 등 KBO 장면 정보를 추출해서 현상황을 정확하게 sLLM 모델이 인지하고 해설을 할 수 있도록 활용할 예정
- 향후 계획: 현재 진행된 박찬호 해설위원 외에 이순철, 김선우 해설위원의 스타일 변환 데이터도 동일한 방식으로 제작하여 파인튜닝을 수행할 예정  
이후 모델별 성능 비교 및 스타일 정확도 평가를 통해 최적의 모델과 학습 전략을 도출할 계획

## 6. 부록

- sLLM 모델 Hugging Face 주소 :  
<https://huggingface.co/SeHee8546/kanana-1.5-8b-pakchanho-lora-v2>
- VLM 모델 Hugging Face 주소 :  
<https://huggingface.co/SeHee8546/qwen3-vl-8b-kbo-scoreboard-qlora-final-V2>