

SK네트웍스 Family AI 과정 12기

## 데이터 전처리 인공지능 학습 결과서

|        |   |
|--------|---|
| 산출물 단계 | 데이터 전처리   |
| 평가 산출물 | 인공지능 학습 결과서   |
| 제출 일자  | 2025.07.20  |
| 깃허브 경로 | <a href="https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN12-FINAL-6TEAM">https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN12-FINAL-6TEAM</a> |
| 작성 팀원  | 이정민, 이지복  |

### 1. 모델 비교 및 선정 이유

- 비교 대상 모델:

| 모델명    | 종류                     | 선정 이유  |
|--------|------------------------|--|
| KoBERT | Transformer 기반 사전학습 모델 | 한국어 문장 의미 파악과 분류에 강한 한국어 사전학습 BERT 기반 모델               |
| KoGPT  | GPT 기반 사전학습 모델         | 자연스러운 한국어 문장 생성에 유리함.<br>대화, 요약, 스토리 작성 등 생성 중심 작업에 적합 |

- 실험 모델 수: 총 2종
- 최종 선정 모델: **KoBERT**

### 2. 모델 구조 및 아키텍처

#### 2.1 모델 아키텍처 도식

- 입력층 → KoBERT Tokenizer → BERT 인코더 블록 → Linear Classifier → Softmax 출력

## 2.2 구성 요소 설명:

| 계층명                 | 역할                               | 구성 요소                            |
|---------------------|----------------------------------|----------------------------------|
| Embedding           | 그림분석 문장을 토큰 벡터로 변환               | 토큰 임베딩, 세그먼트 임베딩, 포지션 임베딩        |
| Encoder             | 입력된 벡터의 문맥 정보를 양방향으로 반영하여 벡터로 변환 | KoBERT 모델 (Multi-Head Attention) |
| Classification Head | 감정 유형 분류 출력                      | Linear Layer + Softmax           |

## 3. 학습 설정 및 하이퍼파라미터

| 항목                  | 값                         |
|---------------------|---------------------------|
| 학습 데이터 수            | 468건                      |
| 검증 데이터 수            | 201건                      |
| 에폭(Epoch) 수         | 15                        |
| 배치 크기 (Batch Size)  | 학습: 8 / 검증: 16            |
| 학습률 (Learning Rate) | 2e-5                      |
| 옵티마이저               | AdamW                     |
| 손실 함수               | CrossEntropyLoss          |
| 평가 지표               | macro_f1                  |
| 조기 종료 기준            | Loss 값이 3회 연속 감소하지 않으면 종료 |

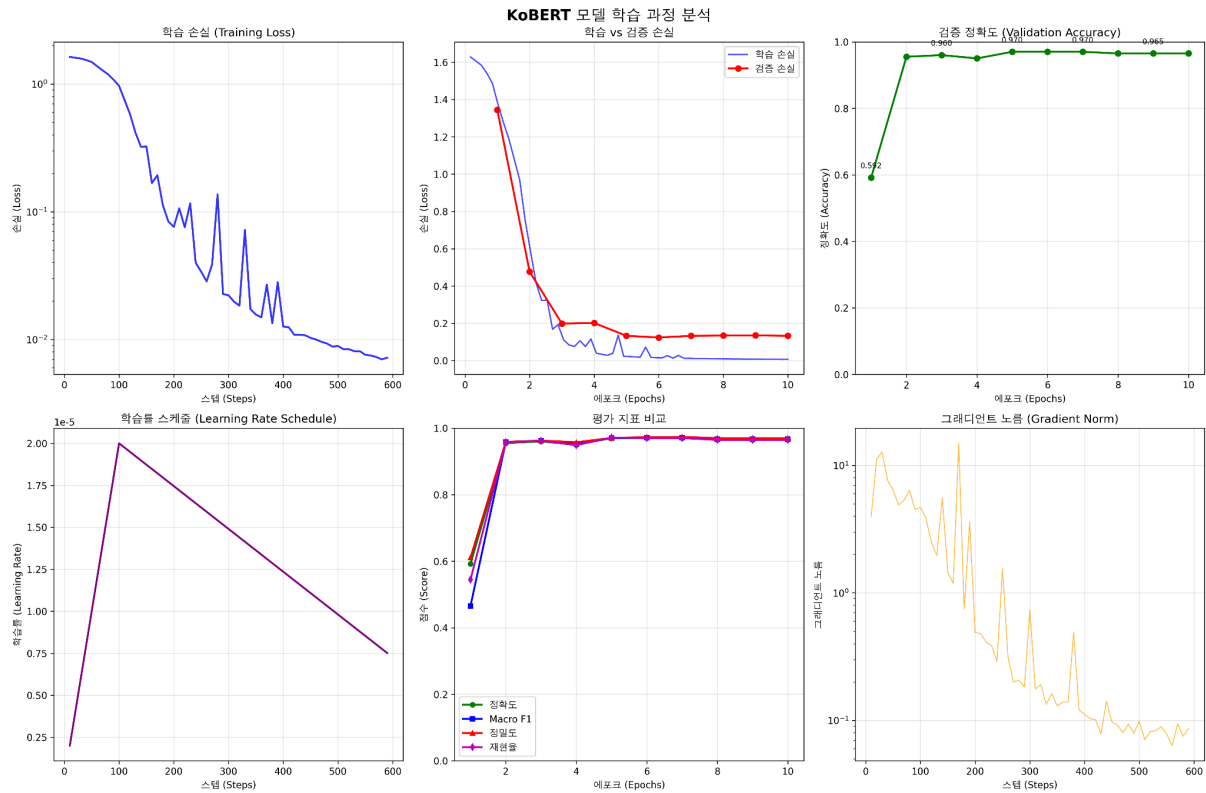
- 학습 데이터는 더 작은 배치로 자주 가중치를 업데이트해서 학습의 안정성을 향상하고, 검증 데이터는 큰 배치로 빠르게 평가하여 시간을 단축함.

## 4. 학습 결과 및 성능 평가

### 4.1. 학습 결과 요약

| 모델     | Accuracy | Precision | Recall | F1 Score |
|--------|----------|-----------|--------|----------|
| KoBERT | 0.965    | 0.97      | 0.96   | 0.967    |
| KoGPT  | 0.940    | 0.95      | 0.94   | 0.943    |

## 4.2. 그래프 : 학습/검증 loss 변화, accuracy 변화 등



## 4.3 해석 및 분석

- **F1 Score가 더 높은 모델: KoBERT**
- **Accuracy, Precision, Recall 등 주요 정량 지표 전반에서 높은 수치를 기록하였으며, 특히 Eval Loss가 낮고 과적합 없이 안정적으로 수렴하는 양상을 보임.** 이는 감정 키워드 기반 문장 분류 과제에 있어 **KoBERT**가 높은 분류 정확도와 신뢰도를 가지는것을 시사함.  
따라서, 모델 자체의 성능을 기반으로 **KoBERT**를 최종 분류 모델로 선정함.
- **향후 검증 예정 :** 본 평가 결과는 고정된 학습/검증 데이터셋을 기준으로 도출된 수치이며, 실제 서비스에서는 **LLM(GPT-4 모델)**이 생성하는 해석 응답이 매 실행마다 표현이 달라질 수 있는 비결정적 특성을 가짐. 이는 감정 유형 분류 모델의 입력 문장에 변동성을 유발하며, 결과적으로 예측 성능의 일관성에도 영향을 미칠 수 있음. 따라서 추후에는 **LLM** 응답을 기반으로 한 테스트 데이터를 **10회** 생성하고, 동일한 분류 모델에 입력하여 얻은 결과의 평균 성능을 기준으로 모델의 강건성과 일관성을 추가 검증할 예정임. 이를 통해 실사용 환경에서의 적용 가능성과 예측 안정성 확보 여부를 종합적으로 판단하고자 함.

## 5. 과적합/과소적합 대응

- 적용 기법:

| 기법      | 설명                       | 적용 여부                        |
|---------|--------------------------|------------------------------|
| Dropout | 과적합 방지를 위해 일부 뉴런 제거      | O (KoBERT 내부에 내장됨)           |
| 조기 종료   | validation 성능 저하 시 학습 중단 | O(Loss 값이 3회 연속 감소하지 않으면 종료) |
| 학습률 감소  | Plateau 시 학습률 자동 감소      | O (ReduceLROnPlateau 사용)     |
| 교차 검증   | 다양한 데이터 분포 반영            | X (단일 validation set 사용)     |

## 6. 결론 및 향후 계획

- 최종 선정 모델: KoBERT
- 활용 방안:
  - 사용자 감정 분석 결과를 5가지 유형(내면형, 안정형, 쾌락형, 관계형, 추진형)으로 분류
  - 분류 결과를 RAG 기반 페르소나 챗봇에 연동
  - 감정 유형별 맞춤형 응답 제공
- 향후 계획:
  - LLM 응답 기반 테스트셋을 10회 반복 실행하여 평균 성능 검증
  - FastAPI 연동 및 REST API 서비스화
  - 모델 경량화 및 실사용 환경 최적화 예정

## 7. 부록

- 학습 코드 경로: train\_and\_eval\_kobert.py
- 데이터 파일: ../backend/data/personality\_keywords\_dataset\_v2.json
- 라벨 매핑: { "추진형": 0, "내면형": 1, "안정형": 2, "관계형": 3, "쾌락형": 4 }
- 모델 및 로그 저장 경로: ./backend/llm/model/kobert\_model/, ./logs/
- Tokenizer: skt/kobert-base-v1
- 전체 학습 로그 캡처 또는 파일: ./backend/llm/model/kobert\_training\_analysis.png
- 학습 코드 경로 또는 GitHub 링크:  
[https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN12-FINAL-6TEAM/blob/backend/backend/llm/model/train\\_and\\_eval\\_kobert.py](https://github.com/SKNETWORKS-FAMILY-AICAMP/SKN12-FINAL-6TEAM/blob/backend/backend/llm/model/train_and_eval_kobert.py)



- 주요 파라미터 설정 config.json

```
{
  "model_name": "skt/kobert-base-v1",
  "architecture": "BertForSequenceClassification",
  "problem_type": "single_label_classification",
  "num_labels": 5,
  "hidden_size": 768,
  "num_hidden_layers": 12,
  "num_attention_heads": 12,
  "hidden_act": "gelu",
  "dropout": 0.1,
  "max_position_embeddings": 512,
  "vocab_size": 8002,
  "initializer_range": 0.02,
  "transformers_version": "4.40.2"
}
```