

LLM파인튜닝의 과정을 도식화

LLM 파인튜닝은 사전 훈련된 대규모 언어 모델을 특정 작업이나 도메인에 최적화하는 과정입니다. 여러 연구와 자료에 따르면 LLM 파인튜닝 과정은 다음과 같이 도식화할 수 있습니다.

LLM 파인튜닝의 7단계 프로세스

1. 데이터 준비 단계

- 데이터 수집 및 정제
- 데이터 포매팅 및 레이블링
- 데이터셋 분할 (훈련/검증/테스트)^{[1] [2]}

2. 모델 초기화 단계

- 사전 훈련된 기본 모델 선택
- 모델 아키텍처 및 크기 결정
- 초기 파라미터 설정^{[1] [3]}

3. 훈련 환경 설정 단계

- 컴퓨팅 리소스 구성
- 필요한 라이브러리 및 프레임워크 설치
- 분산 학습 환경 설정(필요시)^[1]

4. 파인튜닝 실행 단계

- 전체 파인튜닝 또는 부분 파인튜닝 선택
- 하이퍼파라미터 설정 (학습률, 배치 크기, 에폭 수)
- 훈련 실행 및 모니터링^{[1] [4]}

5. 평가 및 검증 단계

- 모델 성능 평가
- 다양한 메트릭을 통한 검증
- 오류 분석 및 개선점 파악^{[1] [5]}

6. 배포 단계

- 최적화된 모델 저장
- 추론 파이프라인 구성
- 서비스 통합^{[1] [6]}

7. 모니터링 및 유지보수 단계

- 성능 모니터링
- 피드백 수집 및 분석
- 필요시 재훈련 및 업데이트^[1]

파인튜닝 방법론

지도 학습 기반 파인튜닝

- 레이블이 있는 데이터셋을 사용
- 입력과 정답 쌍으로 모델 훈련
- 기본 하이퍼파라미터 튜닝 및 전이학습 활용^{[3] [5]}

파라미터 효율적 파인튜닝

- LoRA(Low-Rank Adaptation)
- 어댑터 레이어 활용
- 부분 파인튜닝으로 자원 효율성 증대^{[1] [3]}

파인튜닝 모범 사례

- **명확한 목표 정의:** 구체적인 태스크와 목표를 설정^{[7] [5]}
- **적절한 사전 훈련 모델 선택:** 태스크에 맞는 모델 아키텍처 선정^{[3] [5]}
- **소규모 모델로 시작:** 빠른 실험과 반복을 위해 작은 모델부터 시작^[6]
- **데이터 품질 확보:** 고품질의 대표성 있는 데이터셋 구성^{[5] [6]}
- **하이퍼파라미터 최적화:** 학습률, 배치 크기 등 주요 파라미터 조정^{[5] [6]}
- **다양한 데이터 형식 실험:** 다양한 입력 유형을 포함하여 모델의 범용성 향상^[6]

이러한 단계적 접근을 통해 LLM을 특정 작업에 맞게 최적화하고, 일반 목적 모델을 특화된 솔루션으로 변환할 수 있습니다.^[8]

✻

1. <https://arxiv.org/html/2408.13296v1>
2. <https://hackernoon.com/fine-tuning-an-llm-the-six-step-lifecycle>
3. <https://www.turing.com/resources/finetuning-large-language-models>
4. <https://codelabs.developers.google.com/llm-finetuning-supervised>

5. <https://www.superannotate.com/blog/llm-fine-tuning>
6. <https://www.acorn.io/resources/learning-center/fine-tuning-llm/>
7. <https://www.kdnuggets.com/7-steps-to-mastering-large-language-model-fine-tuning>
8. <https://wandb.ai/byyoung3/Generative-AI/reports/How-to-fine-tune-a-large-language-model-LLM---VmlldzoxMDU2NTg4Mw>