

Prompting 기반 매개변수 효율적인 Few-Shot 학습

박은환*, 서대룡, 전동현, 강인호, 나승훈*

judepark@jbnu.ac.kr*

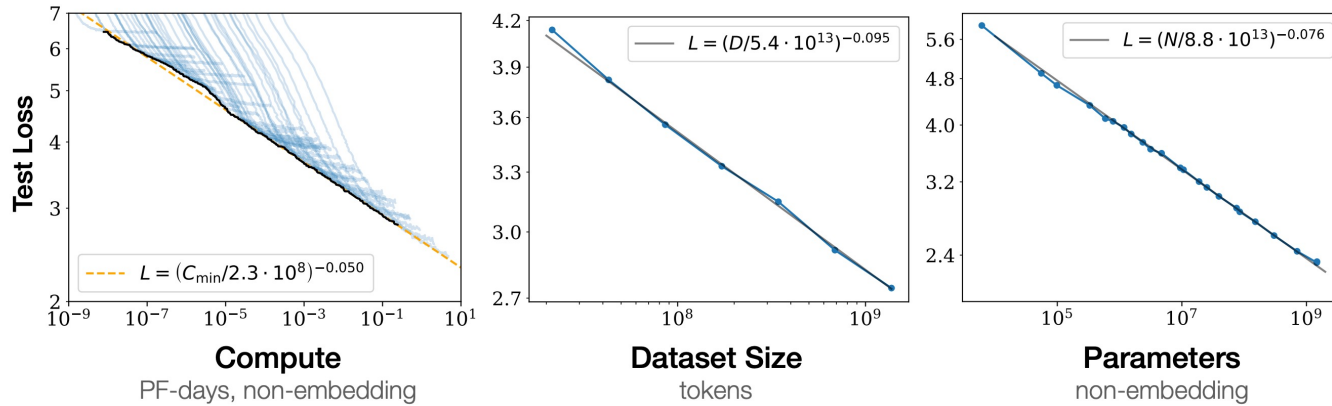


NAVER

Jeonbuk National University*

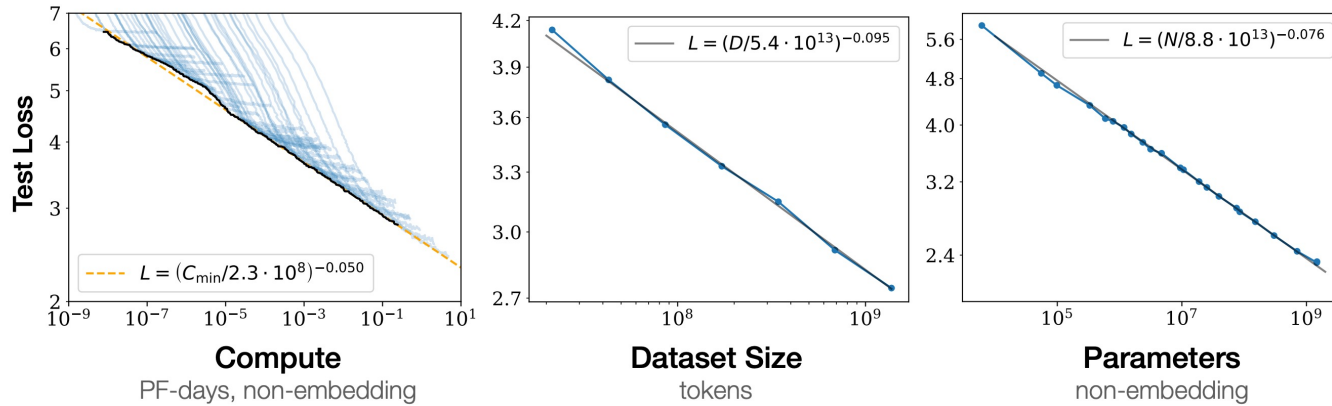
NAVER Corporation

서론



최근 매개 변수와 데이터 집합의 크기 등은 언어 모델의 성능에 많은 영향을 끼치고 있음.

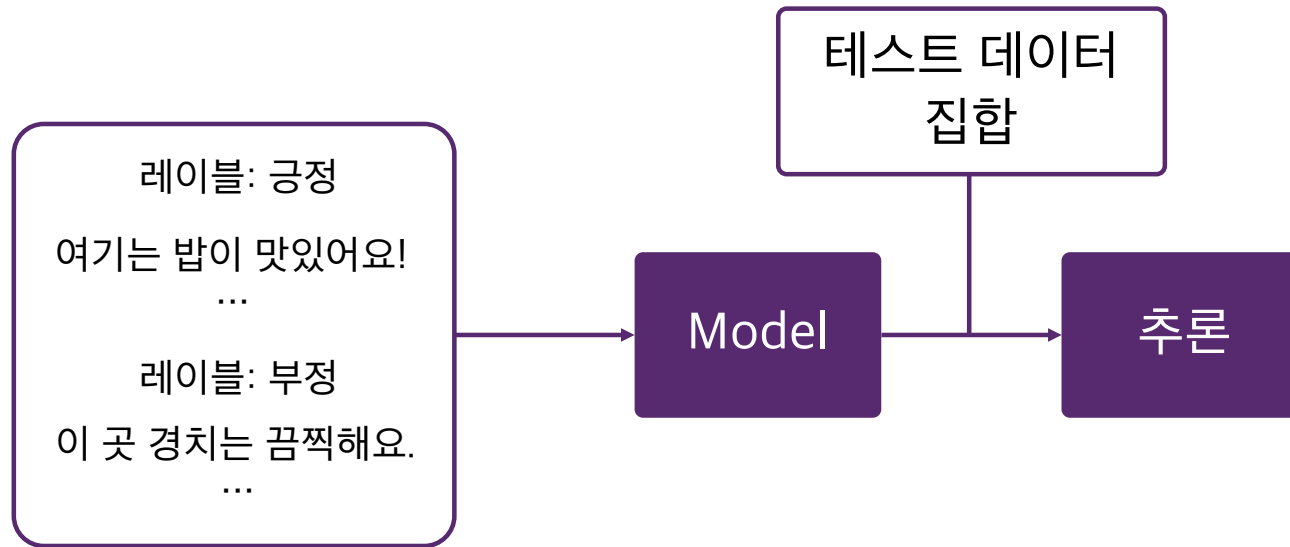
서론



최근 매개 변수와 데이터 집합의 크기 등은 언어 모델의 성능에 많은 영향을 끼치고 있음.

→ 하지만, 방대한 매개 변수의 크기로 인하여 실제 산업에서 쉽게 쓰일 수 없다는 단점이 존재함.

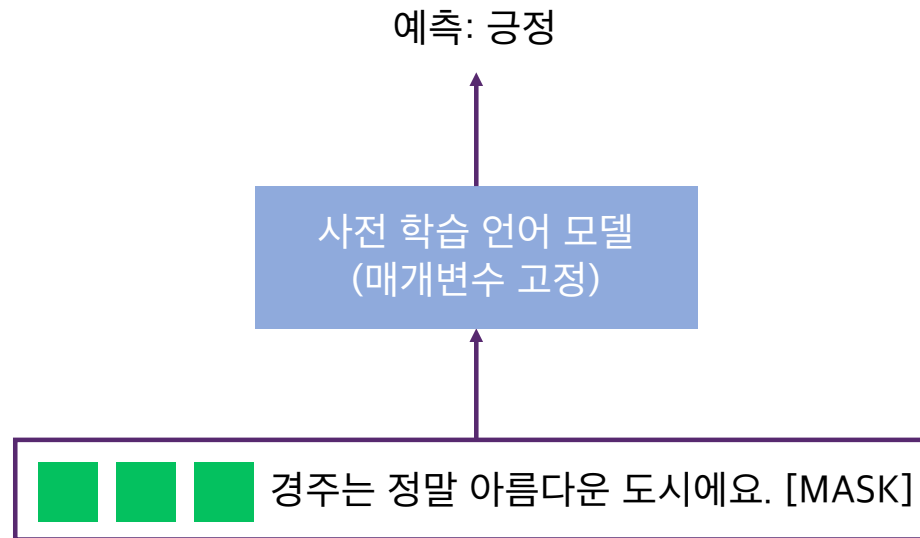
관련 연구 (Few-Shot 학습)



적은 데이터 집합 (예: 레이블 별 K 개)을 구성하여 학습하고 전체 테스트 데이터 집합의 결과를 추론하는 Few-Shot 학습

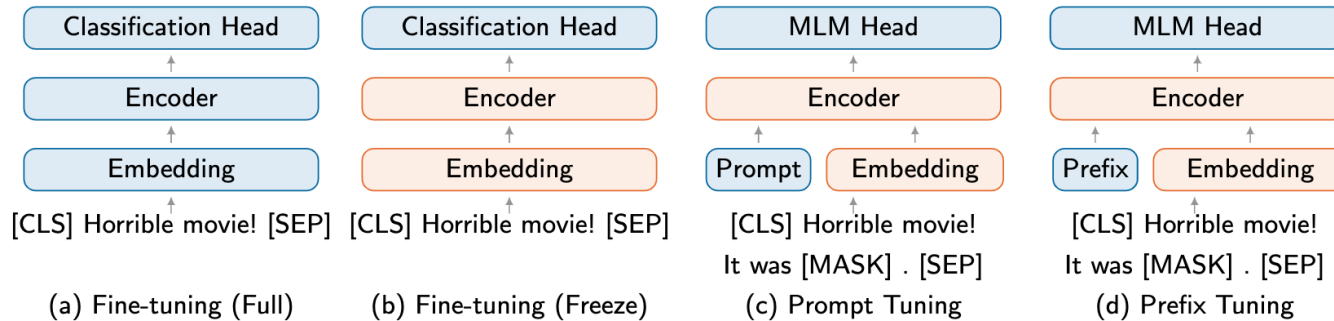
관련 연구 (Prompt-based)

■ 학습 가능한 매개변수



언어 모델의 매개변수는 고정하되, 임베딩 (Soft Prompts) 을 학습하는 Prompt-based Fine-tuning.

제안 방법



- Fine-tuning (Full): 전체 매개변수를 학습하는 일반적인 방식임.
- Fine-tuning (Freeze): Classification Head 만을 학습임.
- Prompt-tuning: Soft Prompt 와 MLM Head 를 학습임.
- Prefix-tuning: Prefix 와 MLM Head 를 학습임.

실험 결과

Model	SST-2 (acc)	RTE (acc)	MR (acc)	CR (acc)
Prompt tuning ($ K = 10$)	88.4 (1.1)	53.5 (2.6)	80.3 (5.1)	88.9 (2.3)
Prompt tuning ($ K = 15$)	88.8 (0.6)	56.1 (1.8)	83.5 (1.1)	87.8 (1.9)
Prompt tuning ($ K = 20$)	88.9 (0.6)	56.0 (0.1)	84.2 (1.1)	89.3 (1.0)
Prefix tuning ($ K = 10$)	81.2 (1.5)	51.7 (0.1)	75.1 (3.8)	83.1 (2.5)
Prefix tuning ($ K = 15$)	80.4 (1.3)	54.1 (0.0)	76.8 (1.9)	83.1 (1.3)
Prefix tuning ($ K = 20$)	81.5 (0.1)	55.9 (0.0)	77.8 (2.1)	84.5 (1.5)
Fine-tuning (Full)	62.2 (8.6)	49.8 (2.5)	53.2 (3.2)	72.8 (6.8)
Fine-tuning (Freeze)	49.0 (0.0)	47.2 (0.0)	50.0 (0.0)	50.0 (0.0)

제안 방법 요약

1. Prompt-based Fine-tuning 을 이용하여 사전 학습된 언어 모델의 매개변수를 학습하지 않을 뿐만 아니라 MLM Head 를 재사용하는 학습 방법을 사용함.
2. Fine-tuning (Full), Fine-tuning (Freeze), Prefix-based fine-tuning, Prompt-based fine-tuning 비교
3. Fine-tuning (Full) 보다 Prefix-, Prompt-based 가 적은 매개변수만을 학습함에도 높은 성능을 보임.

감사합니다.