



LLM, Fine-Tuning

정규 수업 14차시



AI 여친을 구현하기 위한 방법

LLM으로 AI 여친을 구현한다고 가정합시다

LLM으로 AI 여친을 구현한다고 가정합시다

구현할 때에 있어서 조건

1. 사용자를 '자기야' 등으로 불러야함
2. 딱딱한 존댓말을 쓰면 안됨
3. 최신 유행에 대해서 잘 알고 있어야함(20살이라는 설정)

만약 프롬프트 엔지니어링을 사용한다면..

1. 입력이 길어질 수록 모든 In-Context 지식을 반영하기 어려움
2. 입력 토큰이 길기 때문에 답변 속도 낮음, 비용 높음
3. 백사리가 자주남

AI 여친을 구현하기 위한 방법

생각보다 단점이 많죠?

**아무래도 AI 여친을 프롬프트 엔지니어링으로
구현하기에는 어려울 것 같아요!**

지난 시간에 무엇을 했더라?!?!

**지난 시간에 LLM의 한계점을 해결하는 방법으로
프롬프트 엔지니어링을 배웠습니다**

지난 시간에 무엇을 했더라?!?!

하지만....
하나 더 있었죠..?

파인튜닝이란

네~ 바로~

파인튜닝(미세조정, Fine-Tuning)

이었습니다~

파인튜닝이란

파인튜닝(미세조정, Fine-Tuning)

이미 학습을 마친 모델을 추가 데이터를 사용해서 특정 작업이나 도메인에 맞게
미세조정하는 과정을 뜻합니다

여기서 조정이라는 말은 가중치를 조정한다는 것입니다.

파인튜닝이란

파인튜닝이 가지는 최대 장점

적은 비용, 원하는 성능

파인튜닝이란

적은 비용, 원하는 성능

새로운 지식을 알려주기 위해서 프롬프트 엔지니어링처럼 토큰을 소비할 일이 없음

모델의 가중치를 직접 조절함으로써 프롬프트 엔지니어링 보다

훨씬 고수준의 성능을 기대할 수 있음

파인튜닝이란

엥..? 그럼 AI 여친 LLM을 처음 부터 만들면 되지 않나?

LLM을 처음부터 학습 시키는 것은 엄청난 비용과 시간이 소요됨

일반 사용자에게 있어서는 불가능에 가까움

AI에게 말하는 법부터 가르칠 필요가 있을까?

여자친구다운 모습만 알려주면 되는게 아닐까?

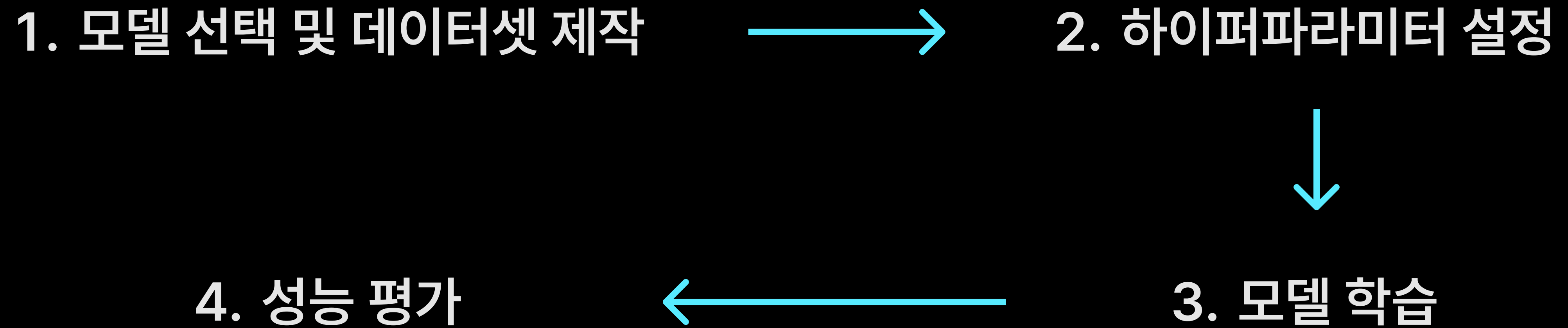
파인튜닝이란

사실 파인튜닝도 비용이 만만치 않습니다.

비용이 적게 든다는 것은.... 동일한 성능을 가지는
파인튜닝된 모델과 프롬프트 엔지니어링한 모델을
단순 사용 시에 파인튜닝 모델이 토큰을 더 적게 사용하여 비용이 적게 든다는 것입니다.

파인튜닝의 취지에 맞게 모델을 처음 만드는 것보다는 비용이 적게 들지만,
그래도 엄청난 시간과 돈이 필요합니다.

파인튜닝의 과정



과정 1 : 모델 선택 및 데이터셋 제작

원하는 데이터를 학습시킬 베이스 모델을 고릅니다.

사용하는 용도와 추구하는 방향에 따라 고를 수 있는 선택지가 여러가지가 됩니다.

고른 베이스 모델에 학습시킬 데이터셋을 제작합니다.

데이터셋을 제작하는 과정에서 데이터를 정제하기 위해서 LLM을 사용하기도 합니다.

과정 2 : 하이퍼파라미터(Hyperparameter) 설정

learning-rate, epochs등 하이퍼파라미터를 설정하여 어느 정도 학습을 할지 정합니다
여기서 설정하는 파라미터에 따라서 같은 데이터셋이라도 다른 성능의 파인튜닝 모델이
나올 수 있습니다.

과정 3 : 모델 학습

이전에 준비한 데이터셋과 설정한 하이퍼파라미터로 모델 학습을 진행합니다.

데이터셋 준비와 함께 가장 시간이 많이 소요되는 작업이며, 하염없이 기다리는 게 일입니다.

이 과정은 수 십 분부터 수 십 시간이 걸릴 수 있습니다.

과정 4 : 성능 평가

학습을 위해서 사용했던 데이터셋으로 학습이 잘 되었나 성능을 평가합니다

성능 평가 후 하이퍼파라미터를 조정하는 등 성능 개선을
위해서 학습을 다시 진행하기도합니다.