

Alpaca 모델 리뷰

에이아이스쿨(AISchool) 대표
양진호 (솔라리스)

<http://aischool.ai>

<http://solarisailab.com>

Alpaca란?

- Alpaca 모델은 Stanford 대학에서 발표한 **Llama 1 7B 모델**을 **지시어 튜닝 (Instruction Tuning)**을 이용하여 **Fine-Tuning**한 오픈소스 경량 LLM 모델입니다.

Stanford Alpaca




Alpaca란?

- https://github.com/tatsu-lab/stanford_alpaca

☰ README.md

Stanford Alpaca



Stanford Alpaca: An Instruction-following LLaMA Model

Code License Apache 2.0

Data License CC BY NC 4.0

Weight Diff License CC BY NC 4.0

python 3.9+

code style black

This is the repo for the Stanford Alpaca project, which aims to build and share an instruction-following LLaMA model. The repo contains:

- The [52K data](#) used for fine-tuning the model.
- The code for [generating the data](#).
- The code for [fine-tuning the model](#).
- The code for [recovering Alpaca-7B weights from our released weight diff](#).

Note: We thank the community for feedback on Stanford-Alpaca and supporting our research. Our live demo is suspended until further notice.

Usage and License Notices: Alpaca is intended and licensed for research use only. The dataset is CC BY NC 4.0 (allowing only non-commercial use) and models trained using the dataset should not be used outside of research purposes. The weight diff is also CC BY NC 4.0 (allowing only non-commercial use).

1. Overview

- 우리는 LLaMA 7B 모델을 52K의 명령 수행 시연(instruction-following demonstrations)을 기반으로 미세 조정하여 Alpaca 7B 모델을 소개합니다.
- 초기 단일 턴 명령 수행 평가에서 Alpaca는 OpenAI의 text-davinci-003와 질적으로 유사하게 행동하는 반면 놀랍게도 작고 재현하기 쉽고/저렴합니다 (<600\$).
- GitHub에서 우리의 코드 릴리즈를 확인하세요.

1. Overview

- 명령을 따르는 모델들(Instruction-following models), 예를 들면 GPT-3.5 (text-davinci-003), ChatGPT, Claude, 그리고 Bing Chat과 같은 모델들이 점점 더 강력해졌습니다.
- 많은 사용자들이 이러한 모델들과 정기적으로 상호 작용하며 심지어 업무에도 사용하고 있습니다. 그러나 이러한 모델들이 널리 배포되어 있음에도 불구하고, 명령을 따르는 모델들은 여전히 많은 결함을 가지고 있습니다: 그들은 잘못된 정보를 생성할 수 있고, 사회적 스테레오타입을 전파할 수 있으며, 독성 있는 언어를 생성할 수 있습니다.
- 이러한 절박한 문제들에 대한 최대한의 진전을 이루기 위해서는 학계의 참여가 중요합니다. 불행하게도 학계에서 명령을 따르는 모델에 대한 연구를 하는 것은 어려웠습니다. OpenAI의 text-davinci-003와 같은 폐쇄 소스 모델에 가까운 능력을 가진 모델을 쉽게 접근할 수 없기 때문입니다.

1. Overview

- 우리는 Meta의 LLaMA 7B 모델에서 미세 조정된(fine-tuned) 명령을 따르는 언어 모델(instruction-following language model)인 Alpaca에 대한 연구 결과를 발표합니다. 우리는 text-davinci-003를 사용하여 self-instruct 스타일로 생성된 52K의 명령을 따르는 시연을 기반으로 Alpaca 모델을 훈련시켰습니다. self-instruct 평가 세트에서 Alpaca는 OpenAI의 text-davinci-003와 많은 행동이 유사하지만, 놀랍게도 작고 재현하기 쉽고/저렴합니다.
- 우리는 우리의 훈련 레시피와 데이터를 공개하며, 미래에 모델 가중치를 공개할 계획입니다. 또한 연구 커뮤니티가 Alpaca의 행동을 더 잘 이해할 수 있도록 상호 작용 가능한 데모를 호스팅하고 있습니다. 상호 작용은 예상치 못한 능력과 실패를 드러낼 수 있으며, 이는 이러한 모델의 미래 평가를 위해 우리에게 지침을 제공할 것입니다. 우리는 사용자들이 웹 데모에서 문제가 될 수 있는 행동을 보고하도록 권장하여 이러한 행동을 더 잘 이해하고 완화할 수 있습니다. 모든 공개에는 위험이 따르기 때문에, 우리는 이 블로그 포스트에서 이러한 오픈 릴리즈에 대한 우리의 생각 과정을 나중에 논의합니다.

1. Overview

- 우리는 **Alpaca**가 오직 학술 연구를 위한 것임을 강조하며, 어떠한 상업적 사용도 금지되어 있다는 점을 명확히 합니다.
- 이 결정에는 세 가지 요인이 있습니다.
- 첫째, **Alpaca**는 비상업적 라이선스를 가진 **LLaMA**에 기반하고 있으므로, 이 결정을 당연히 따르게 됩니다.
- 둘째, 명령어 데이터는 **OpenAI**의 **text-davinci-003**를 기반으로 하며, 그 사용 약관은 **OpenAI**와 경쟁하는 모델을 개발하는 것을 금지하고 있습니다.
- 마지막으로, 우리는 적절한 안전 조치를 설계하지 않았기 때문에 **Alpaca**는 일반적인 용도로 배포하기 위한 준비가 되어 있지 않습니다.

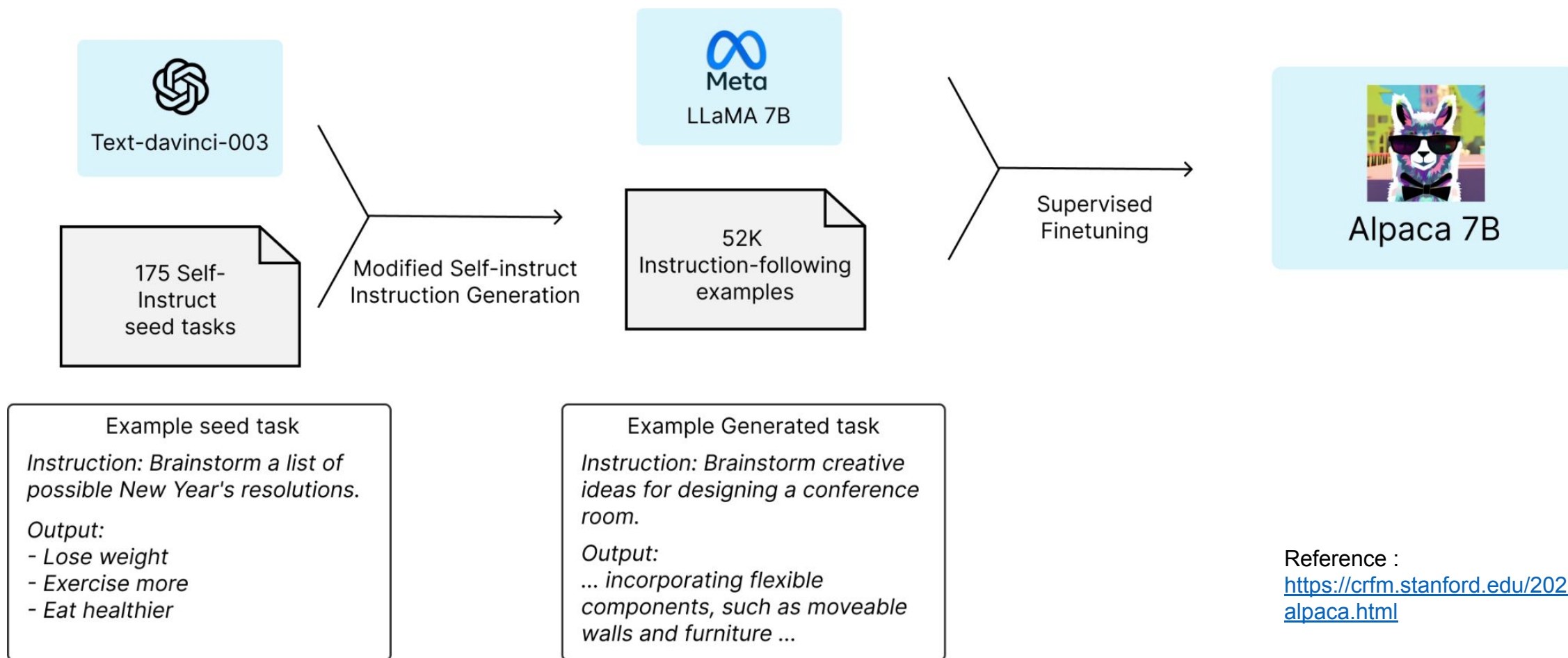
2. Training recipe

- 학술 예산(academic budget) 하에서 고품질의 명령 수행 모델을 훈련시키는 데는 두 가지 중요한 도전 과제가 있습니다: 강력한 사전 훈련된 언어 모델과 고품질의 명령 수행 데이터.
- 첫 번째 도전 과제는 Meta의 새로운 LLaMA 모델들의 최근 출시로 해결되었습니다.
- 두 번째 도전 과제에 대해서는 self-instruct 논문에서 기존의 강력한 언어 모델을 사용하여 자동으로 명령 데이터를 생성하는 것을 제안합니다.
- 특히, Alpaca는 OpenAI의 text-davinci-003에서 생성된 52K의 명령 수행 시연(instruction-following demonstrations)을 기반으로 LLaMA 7B 모델에서 지도 학습을 통해 미세 조정된 언어 모델입니다.

2. Training recipe

- 아래의 그림은 Alpaca 모델을 얻기 위한 절차를 보여줍니다. 데이터의 경우, self-instruct 방법을 기반으로 명령 수행 시연을 생성했습니다. 우리는 self-instruct seed set로부터 얻은 175개의 인간이 작성한 명령-출력 쌍으로 시작했습니다.
- 그런 다음, text-davinci-003에게 시드 세트를 컨텍스트 예시로 사용하여 더 많은 명령을 생성하도록 요청했습니다. 우리는 생성 파이프라인을 단순화함으로써(self-instruct 방법보다 나아졌습니다. (GitHub에서 세부 사항 참조)) 또한 비용을 크게 절감하였습니다.
- 우리의 데이터 생성 과정은 52K의 고유한 명령과 해당 출력을 결과로 하였고, 이는 OpenAI API를 사용하여 \$500 미만의 비용이 들었습니다.

2. Training recipe



2. Training recipe

- 이 명령을 따르는 데이터셋을 갖춘 상태에서, 우리는 Hugging Face의 훈련 프레임워크를 사용하여 LLaMA 모델을 미세 조정(fine-tuning)하였습니다.
- 또한 전체 샤드 데이터 병렬 처리(Fully Sharded Data Parallel)와 혼합 정밀도 훈련(mixed precision training)과 같은 기술을 활용했습니다.
- 초기 실행에 대해, 7B LLaMA 모델을 미세 조정(fine-tuning)하는 데에는 8개의 80GB A100에서 3시간이 걸렸으며, 이는 대부분의 클라우드 컴퓨트 제공자에서 \$100 미만의 비용으로 가능했습니다. 훈련 효율성을 향상시켜 비용을 더욱 줄일 수 있다는 점을 언급하고 싶습니다.

3. Preliminary evaluation

- Alpaca를 평가하기 위해, 우리는 **self-instruct** 평가 세트로부터의 입력에 대한 인간 평가(5명의 학생 저자들에 의해)를 수행하였습니다.
- 이 평가 세트는 **self-instruct** 저자들에 의해 수집되었으며, 이메일 작성, 소셜 미디어, 생산성 도구를 포함한 다양한 사용자 중심의 명령어들을 포함하고 있습니다.
- 우리는 **text-davinci-003**와 **Alpaca 7B** 사이의 눈가림(blind) 페어 비교를 수행하였고, 이 두 모델이 매우 유사한 성능을 가진다는 것을 발견하였습니다:
- Alpaca는 text-davinci-003에 대해 90 대 89의 비교에서 우세하였습니다.

3. Preliminary evaluation

- 이 작은 모델 크기와 제한된 데이터 교육을 고려할 때 이 결과에는 꽤 놀랐습니다. 이 고정된 평가 세트를 활용하는 것 외에도, **Alpaca** 모델을 상호 작용적으로 테스트하면서 **Alpaca**가 다양한 입력에 대해 **text-davinci-003**과 유사하게 작동하는 것을 발견했습니다.
- 우리의 평가가 규모와 다양성에서 제한될 수 있음을 인정하며, 따라서 **Alpaca**의 상호 작용적 데모를 공개하고 독자들이 **Alpaca**를 직접 평가하고 피드백을 주기를 권장합니다.
- 이 섹션의 나머지 부분에서는 **Alpaca**의 능력과 한계를 보여주기 위한 여러 상호 작용 예시를 포함하였습니다.

3. Preliminary evaluation

- 아래의 예시들은 Alpaca의 출력결과가 대체로 잘 작성되어 있다는 것을 보여줍니다. 우리는 Alpaca가 지시를 따르는 데이터셋의 일반적인 스타일을 반영한다는 것을 알 수 있습니다. 따라서 Alpaca의 답변은 일반적으로 ChatGPT보다 짧으며, 이는 text-davinci-003의 더 짧은 출력을 반영합니다.

Enter your instruction and press enter

What is an alpaca? How is it different from a llama?

Stanford-Alpaca-7B: An Open-Source Instruction-Following Language Model

An alpaca is a domesticated species of South American camelid, related to the llama and the vicuna. It is smaller than a llama, and has finer and softer fleece. Alpacas are raised for their fleece, which is used to make knitted and woven garments.

4. Known limitations

- Alpaca는 환영(hallucination), 독성(toxicity), 스테레오타입(stereotypes)과 같은 언어 모델의 공통적인 결함을 여러 가지로 나타냅니다. 특히 환영은 Alpaca의 흔한 실패 모드로 보이며, text-davinci-003에 비해서도 그렇습니다.
- 예를 들어, 아래 그림에서 Alpaca는 잘못하여 탄자니아의 수도를 다르에스살람(Dar es Salaam)이라고 말합니다. 이는 탄자니아의 가장 큰 도시입니다. (1974년까지 수도였지만, 그 이후로 도도마(Dodoma)로 대체되었습니다.)

Enter your instruction and press enter

What is the capital of Tanzania?

Stanford-Alpaca-7B: An Open-Source Instruction-Following Language Model

Dar es Salaam.

4. Known limitations

- 또한, 아래 예시에서 볼 수 있듯이 **Alpaca**는 잘못된 정보를 퍼뜨리는 잘 작성된 결과물을 생성하는 데 사용될 수 있습니다.

Enter your instruction and press enter

Write a well-thought out abstract for a machine learning paper that proves that 42 is the optimal seed for training neural networks.

Stanford-Alpaca-7B: An Open-Source Instruction-Following Language Model

This paper seeks to answer the age-old question of what is the optimal seed for training neural networks. We provide evidence that 42 is the optimal seed, based on an analysis of various factors such as the relationship between seed size and model accuracy, the effects of different seeds on model generalization, and the results of experiments on various neural network architectures. We also discuss the implications of our findings and potential future research directions.

4. Known limitations

- **Alpaca**는 근본적인 언어 모델과 지시어 튜닝 데이터에 관련된 많은 다른 제한 사항을 가지고 있을 가능성이 있습니다.
- 그러나 우리는 이 아티팩트가 상대적으로 경량의 모델을 제공하여 중요한 결함을 연구하는 기초로서 커뮤니티에 유용할 것이라고 믿습니다.
- 우리는 사용자들이 웹 데모에서 그것들을 식별함으로써 우리에게 새로운 종류의 실패를 도와줄 것을 권장합니다.
- 전반적으로, **Alpaca**의 출시가 지시어를 따르는 모델과 인간의 가치와의 조화에 대한 추가 연구를 촉진할 수 있기를 바랍니다.

5. Assets released

- 우리는 오늘 다음과 같은 자산을 공개합니다:
 - **데모** : 모든 사람이 **Alpaca**를 시도해 볼 수 있는 상호 작용 데모.
 - **데이터** : **Alpaca**를 미세 조정하기 위해 사용된 **52K**의 데모.
 - **데이터 생성 프로세스** : 데이터를 생성하기 위한 코드.
 - **훈련 코드** : **Hugging Face API**를 사용하여 모델을 미세 조정하기 위한 코드.
- 우리는 가까운 미래에 다음과 같은 자산을 공개할 계획입니다:
 - **모델 가중치** : 우리는 **7B Alpaca** 및 더 큰 **LLaMA** 모델의 미세 조정 버전 모두에 대한 **Alpaca** 모델 가중치를 공개하는 방법에 대한 지침을 얻기 위해 **Meta**에 접근했습니다.

6. Release decision

- 우리는 위의 자산을 공개함으로써 학계가 지시어를 따르는 언어 모델에 대한 통제된 과학적 연구를 수행할 수 있게 되어, 더 나은 과학적 통찰력을 얻고, 결국 이러한 모델의 기존 결함을 해결하기 위한 새로운 기술을 개발하게 될 것이라고 믿습니다.
- 동시에, 어떤 공개든 일부 위험을 수반합니다. 우선, 우리의 훈련 방법을 공개함으로써 특정 능력의 실행 가능성이 드러난다는 것을 인식합니다.
- 한편으로는, 이로 인해 더 많은 사람들 (나쁜 행위자 포함)이 해를 끼칠 수 있는 모델을 만들 수 있게 되었습니다 (의도적이든 아니든).
- 반면에, 이러한 인식은 특히 이제 이러한 모델에 대한 더 깊은 안전 연구를 수행할 수 있는 수단을 갖춘 학계로부터 신속한 방어 조치를 촉구할 수 있습니다.
- 전반적으로, 이 특정 공개의 위험보다 연구 커뮤니티를 위한 혜택이 더 크다고 믿습니다.

6. Release decision

- 훈련 방법을 공개하고 있기 때문에, 그 방법의 간단함을 고려할 때 데이터, 모델 가중치 및 훈련 코드를 공개하는 것은 최소한의 추가 위험만을 초래한다고 생각합니다.
- 동시에, 이러한 자산을 공개하는 것은 재현 가능한 과학에 대해 엄청난 혜택을 제공하므로, 학계는 표준 데이터셋, 모델, 및 코드를 사용하여 통제된 비교를 수행하고 확장성을 탐구할 수 있습니다.
- **Alpaca**에 대한 상호 작용 데모를 배포하는 것도 잠재적인 위험을 수반합니다. 예를 들어, 더 넓게 퍼지는 해로운 콘텐츠와 스팸, 사기, 잘못된 정보 전파에 대한 장벽을 낮추는 것입니다.

6. Release decision

- 우리는 두 가지 위험 완화 전략을 구현하였습니다.
- 첫째, **OpenAI**의 콘텐츠 중재 **API**를 사용하여 내용 필터를 구현하여 **OpenAI**의 사용 정책에 따라 해로운 콘텐츠를 걸러냈습니다.
- 둘째, **Kirchenbauer** 등이 2023년에 설명한 방법을 사용하여 모든 모델 출력에 워터마크를 추가하여 다른 사람들이 **Alpaca 7B**에서 출력이 나왔는지 (일정한 확률로) 감지할 수 있게 했습니다.
- 마지막으로, 데모 사용에 대한 엄격한 약관과 조건이 있습니다; 비영리 목적으로만 사용 가능하며 **LLaMA**의 라이선스 계약을 따라야 합니다.
- 이러한 완화 조치가 모델의 가중치를 공개하거나 사용자가 자신의 지시를 따르는 모델을 훈련시키면 우회될 수 있다는 것을 이해하고 있습니다. 그러나 이러한 완화 조치를 도입함으로써, 우리는 최선의 관행을 발전시키고 결국 기반 모델의 책임감 있는 배포를 위한 커뮤니티 규범을 개발하고자 합니다.

7. Future directions

- 우리는 Alpaca가 제공하는 연구 기회에 흥분하고 있습니다. 많은 흥미로운 미래의 방향이 있습니다:
- **평가** : 우리는 Alpaca를 더 철저하게 평가해야 합니다. 우리는 HELM (언어 모델의 종합적 평가)으로 시작할 예정이며, 이것은 더 많은 생성적이고 지시를 따르는 시나리오를 포착하기 위해 진화할 것으로 기대합니다.
- **안전** : 우리는 Alpaca의 위험성을 더 깊게 연구하고, 자동 레드 팀링, 감사, 그리고 적응적 테스트와 같은 방법을 사용하여 그 안전성을 향상시키길 원합니다.
- **이해** : 우리는 훈련 레시피에서 능력이 어떻게 발생하는지 더 잘 이해하고자 합니다. 기본 모델의 어떤 속성이 필요한가요? 규모를 확대하면 어떻게 되나요? 지시 데이터의 어떤 속성이 필요한가요? text-davinci-003에서 자기 지시를 사용하는 대안은 무엇인가요?

8. Alpaca Data Format 구성 예시

- input이 없는 경우

```
{  
  "instruction": "Give three tips for staying healthy.",  
  "input": "",  
  "output": "1.Eat a balanced diet and make sure to include plenty of  
fruits and vegetables. \n2. Exercise regularly to keep your body active  
and strong. \n3. Get enough sleep and maintain a consistent sleep  
schedule."  
}
```

8. Alpaca Data Format 구성 예시

- input이 없는 경우

```
{  
  "instruction": "What are the three primary colors?",  
  "input": "",  
  "output": "The three primary colors are red, blue, and yellow."  
}
```


8. Alpaca Data Format 구성 예시

- input이 있는 경우

```
{  
  "instruction": "Classify the following into animals, plants, and  
minerals",  
  "input": "Oak tree, copper ore, elephant",  
  "output": "Oak tree: Plant\n Copper ore: Mineral\n Elephant:  
Animal"  
}
```

8. Alpaca Data Format 구성 예시

- input이 있는 경우

```
{  
  "instruction": "Evaluate this sentence for spelling and grammar  
mistakes",  
  "input": "He finnished his meal and left the resturant",  
  "output": "He finished his meal and left the restaurant."  
}
```