MiniCPM-QA

Qi He
22307130442@m.fudan.edu.cn
School of Data Science, Fudan University, China

Abstract

One effective strategy to enhance LLM performance is instruction fine-tuning. In this research, I fine-tune MiniCPM, a compact LLM variant, on two similar datasets to validate the efficacy of instruction-based fine-tuning. Then I evaluate the fine-tuned model on a separate dataset to assess its robustness across diverse tasks. This study underscores the potential of instruction learning to enhance LLM adaptability and performance in various practical applications.

1 Introduction

Proven machine learning algorithms and advanced graphics processing unit breed the rapid development of LLM (large language model). With the fierce springing of large models, people have focused on modifying models on specific tasks. Researchers have proved instruction learning is an efficient way of training method for large language models. It improves models' performance by adding an instructive prefix like 'answer the following questions according to texts'. Instruction learning can also improve the ability of generation, performing better on zero-shot and few-shot tasks. In this research, a small type of LLM, MiniCPM, is fine-tuned on two similar datasets to prove the efficiency of instruction fine-tuning process. By comparing prediction accuracy before and after training, we come out with effect of instruction fine-tuning in the aspect of understanding context and modifying answers.
Afterwards, the tuned model is trained on the other dataset and repeats predicting answers for the former one to test the robustness of fine-tuning process.

2 Related Work

Recent studies have highlighted several advancements in improving the capabilities of large language models (LLMs) through innovative approaches:

2.1 Instruction compliance for LLMs

Task-Agnostic Prefix Prompt (TAPP): Ye et al. (2023) demonstrated that adding task-agnostic prefix prompts to inputs enhances LLMs' ability to follow instructions effectively.
Self-instructed Models: Wang et al. (2023) proposed self-instructed models where the LLM generates its own instructions due to the limited availability of human-generated instructions.
Chain of Thoughts: Recognizing the complexity of math problems, instruction learning has been extended to aid LLMs in solving mathematical tasks. For instance, Wei et al. (2023) introduced the 'Chain-of-Thoughts' method, breaking down problems into smaller, more manageable subproblems for LLMs.
Scaling Instruction Fine-tuning: Researchers have found that scaling up the number of tasks and sizes of models can significantly enhance the effectiveness of instruction fine-tuning processes. These advancements underscore ongoing efforts to leverage instruction learning and related techniques to expand the capabilities of LLMs across various domains, from following instructions to solving complex mathematical problems.

2.2 Prompt engineering

Prompt learning, akin to instruction learning, is a method of fine-tuning that leverages cues for pre-trained language models to enhance their comprehension of human queries. As described by Pengfei Liu, a prompt serves as guidance for the model, aiding in its ability to grasp the nuances of questions posed by humans. It involves enriching the input by incorporating supplementary text, thereby optimizing the utilization of the pre-existing knowledge within these models. In the age of pre-trained and fine-tuned models, devising an effective prompt learning approach stands as pivotal in elevating the performance of these expansive models

2.3 Domain-oriented instruction data set construction

DISC-LawLLM is an intelligent legal system based on large language models to provide a wide range of legal services. It includes case analysis, legal consultation, statute retrieval, and other legal-related services. The construction of DISC-Law-SFT Datasets combines legal task datasets, legal raw text, and open-source instruction datasets as data sources.

DISC-MedLLM provides conversational healthcare services using finetuned large language models. It collects data from real-world diagnosing dialogues, medical knowledge graphs, and manually selected high-quality dialogues. After human and LLM reconstruction, the original data is processed into the DISC-Med-SFT Dataset. Despite potentially lower accuracy compared to GPT-3.5, DISC-MedLLM excels in proactivity and helpfulness.

DISC-FinLLM stands as an intelligent finance system handling financial QA, information extraction, financial calculations, investment advice, and current affairs analysis. DISC-Fin-SFT Datasets consist of financial consulting instructions, financial task instructions, financial computing instructions, and retrieval-enhanced instructions. During QA pair data construction, ChatGPT is used to generate related data. DISC-FinLLM comprises four LoRA modules trained from the base model, ensuring each module specializes in distinct financial domains.

SoMeLVLM is a Large Vision Language Model for Social Media Processing, possessing capabilities across knowledge, comprehension, application, analysis, evaluation, and creation. It collects extensive datasets covering emotions, humor, social factors, etc. The structure of SoMeLVLM follows the cognitive pyramid and is finetuned with instruction-tuning data across all five cognitive levels.

2.4 Supervised learning and efficient fine-tuning strategy

For a pre-trained base model, the original weight matrix W is kept unchanged, and only the training update part is fine-tuned, and the update weight matrix is decomposed into two low-rank matrices A and B. It sharply cuts down the number of trainable parameter by 10,000 times and saves GPU memory requirement by 3 times.

3 Datasets

The dataset for MiniCPM-QA consists of two main high-quality datasets. They share the same key features, which include the question-answer pairs and the context or evidence demonstrating the output answer of each data.

Cosmos QA, presented at EMNLP'2019, is designed for Machine Reading Comprehension with Contextual Commonsense Reasoning. It consists of 35.6K problems in the form of multiple-choice questions that require understanding based on commonsense reasoning. The dataset challenges comprehension beyond literal text spans, focusing on interpreting people's everyday narratives to reason about causes or effects of events.

TriviaQA is a reading comprehension dataset comprising more than 650,000 question-answer-evidence triples. It features 95,000 question-answer pairs created by trivia enthusiasts, accompanied by independently collected evidence documents—six on average per question—that offer high-quality distant supervision for answering the questions. Further details can be found in this paper.

Regarding file inclusion:

1. `qa/wikipedia-train.json`,
   `qa/web-train.json`
   (train dataset)

2. `qa/[verified-]wikipedia-dev.json`,
   `qa/[verified-]web-dev.json`
   (test/validation dataset)

3. `qa/wikipedia-test-without-answers.json`,
   `qa/web-test-without-answers.json`

Directories:

- `evidence/web`
- `evidence/Wikipedia`

For the sake of better training effects, `qa/wikipedia-test-without-answers.json` and `qa/web-test-without-answers.json` are not utilized because they cannot contribute to the training process.

4 Experiment

The pretrained model is MiniCPM-2B-sft-fp32, from ModelBest Inc. and TsinghuaNLP, which contains only 2.4B parameters excluding embeddings.

4.1 CosmosQA

Baseline model: The raw data is processed by adding the role for different content to help the pre-trained model learn information from Multi-round dialogue.

Input Example

```
<user>Choose the correct option
according to the following
context | Question:|Option0-3:
Following is the context.
"Context"
<assistant>"Option_1."
```

After directly preprocessing the raw data and give it to the pretrained MiniCPM, it produces the answer option with some simple explanations.

Output Example

```
Option_1: It was a life or death
emergency situation.
```

Before calculating the accuracy of the baseline predictions, the resulting data should undergo a regular expression process to meet the evaluation standard.

Introduction fine-tuned model: Considering the amount of data in the training file, the training epoch is set as 1 to prevent overfitting. Post-training, the model precisely learns how to provide answers as required by the dataset without additional demonstrations, producing the exact expected option.

Output Example

```
Option_1.
```

Analysis

Model	Accuracy
Baseline	0.5281
Fine-tuned	0.7313

Table 1: CosmosQA Acurracy

The significant improvement achieved through instruction fine-tuning demonstrates that this training method effectively helps the pre-trained model identify essential information in lengthy articles and distinguish key differences between options. Additionally, instruction fine-tuning contributes to producing more regular answers, reducing the effort required to extract the expected information from responses

4.2 TriviaQA

Comparing to the Cosmos, the QA difficulty of Trivia dataset is apparently higher because it not only command answers filling instead of choosing, but provides a longer and diverse context, putting forward a challenging problem for large language model to understand articles and search answer information.

Baseline model:feed the pretrained model with preprocessed instruction fine-tuned data and invoke functions of chatting to generate model predictions.

Input Example

```
<user>Answer the question
according to the following
context | Question:
Following is the context.
"Context"
<assistant>"Answer."
```

Output Example

```
The title song for the Bond
film Octopussy was sung by Rita
Coolidge.
```

Instruction fine-tuned model: Train the pretrained MiniCPM on the preprocessed dataset, then use the tuned model to predict answers. Due to the lengthy context in TriviaQA, especially the Wikipedia section, and limited GPU memory, the maximum token length is restricted to 1024. Additionally, during prediction, the maximum number of input tokens is set to 8000 to accelerate the generation process.

Output Example

```
Octopussy.
```

Interfered model: Select the CosmosQA dataset as the interfering dataset. Train the fine-tuned model using the preprocessed CosmosQA dataset, and then make predictions using the trained model

Output Example

```
The title song for the Bond
film Octopussy was sung by Rita
Coolidge.
```

Analysis The pre-trained model initially performs poorly on both Web and Wikipedia question-answer problems, often failing to provide correct answers. However, upon closer examination of

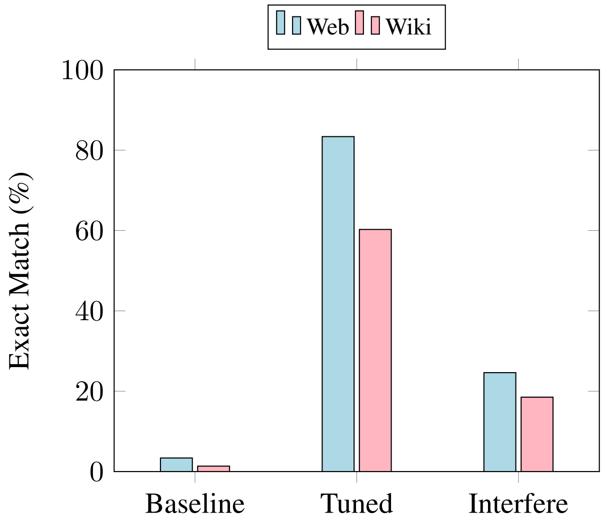

Figure 1: Exact Match Performance of Web and Wiki Models

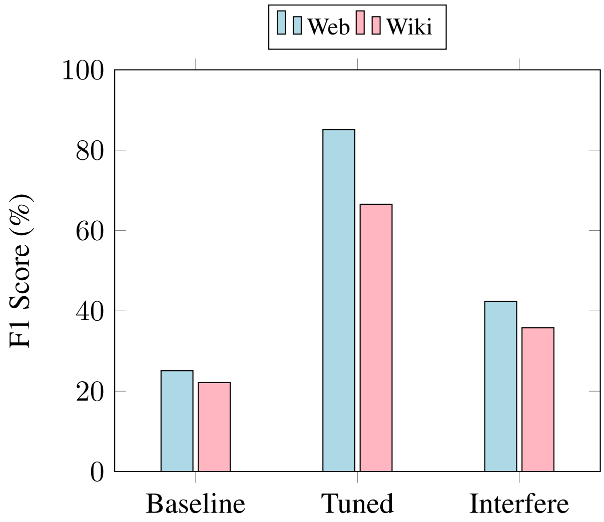

Figure 2: F1 Score Performance of Web and Wiki Models

its incorrect responses, we observe that while they do not meet the standard for correctness, they often contain fragments or elements of the correct answer. This phenomenon suggests that the pre-trained model possesses the capability to comprehend the articles and generate relevant responses.
In contrast, the fine-tuned model excels particularly on the web dataset, achieving an F1 score of over 85. This demonstrates the tuned model's ability to accurately generate answers based on provided text. The lower performance of the fine-tuned model on the Wikipedia dataset may be attributed to the extensive length of context. This not only complicates context comprehension but also leads to incomplete text inputs during both training and prediction processes. Overall, fine-tuning proves to be an effective method for adapting a pre-trained model to specific tasks.
However, interference from training on other datasets can significantly diminish the performance of fine-tuned models on targeted tasks. As illustrated in the chart, the interfered model scores only 42.35 and 35.79 in F1 grade on the two test datasets. This highlights the fine-tuned model's sensitivity and specificity, suggesting that improving the overall performance of large language models across diverse tasks cannot solely rely on fine-tuning based on a few datasets.

5 Conclusion

Instruction fine-tuning has proven effective across both datasets by standardizing answer formats, enhancing article comprehension, and identifying options, resulting in a minimum 20 percent increase in test dataset accuracy.
However, the fine-tuned model exhibits limited and fragile generalization ability. Enhancing the capabilities of large language models across diverse natural language tasks may require additional training parameters or novel training methodologies.
As outlined in the introduction, this experiment adopted multiple-turn dialogue as the introduction method, with further exploration needed to discover more effective introduction strategies.

References

Zhijie Bao, Wei Chen, Shengze Xiao, Kuang Ren, Jiaao Wu, Cheng Zhong, Jiajie Peng, Xuanjing Huang, and Zhongyu Wei. 2023. Disc-medllm: Bridging general large language models and real-world medical consultation.

Wei Chen, Qiushi Wang, Zefei Long, Xianyin Zhang, Zhongtian Lu, Bingxuan Li, Siyuan Wang, Jiarong Xu, Xiang Bai, Xuanjing Huang, and Zhongyu Wei. 2023. Disc-finllm: A chinese financial large language model based on multiple experts fine-tuning.

Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Alex Castro-Ros, Marie Pellat, Kevin Robinson, Dasha Valter, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. 2022. Scaling instruction-finetuned language models.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. Lora: Low-rank adaptation of large language models.

Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. 2023. Self-instruct: Aligning language models with self-generated instructions.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. 2023. Chain-of-thought prompting elicits reasoning in large language models.

Seonghyeon Ye, Hyeonbin Hwang, Sohee Yang, Hyeongu Yun, Yireun Kim, and Minjoon Seo. 2023. Investigating the effectiveness of task-agnostic prefix prompt for instruction following.

Shengbin Yue, Wei Chen, Siyuan Wang, Bingxuan Li, Chenchen Shen, Shujun Liu, Yuxuan Zhou, Yao Xiao, Song Yun, Xuanjing Huang, and Zhongyu Wei. 2023. Disc-lawllm: Fine-tuning large language models for intelligent legal services.

Xinnong Zhang, Haoyu Kuang, Xinyi Mou, Hanjia Lyu, Kun Wu, Siming Chen, Jiebo Luo, Xuanjing Huang, and Zhongyu Wei. 2024. Somelvlm: A large vision language model for social media processing.

Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, Yifan Du, Chen Yang, Yushuo Chen, Zhipeng Chen, Jinhao Jiang, Ruiyang Ren, Yifan Li, Xinyu Tang, Zikang Liu, Peiyu Liu, Jian-Yun Nie, and Ji-Rong Wen. 2023. A survey of large language models.