

Coconut

Description

Coconut¹

Abstract: This paper introduces a new framework for training Large Language Models (LLMs) to reason in a continuous latent space. The framework, named Coconut, leverages a novel architecture that combines a pre-trained LLM with a custom-designed reasoning module. This module enables the model to generate coherent and logical reasoning paths for complex tasks, significantly improving its performance on benchmarks that require deep understanding and logical deduction. The proposed method is evaluated on various tasks, demonstrating superior results compared to existing baselines.

arXiv:2405.14114v1 [cs.LG] 2024-05-20
Training Large Language Models to Reason in a Continuous Latent Space
Coconut

LLM² is a powerful tool for generating text, but it often struggles with reasoning tasks. Coconut addresses this by introducing a reasoning module that allows the LLM to generate coherent and logical reasoning paths. This module is trained using a novel architecture that combines a pre-trained LLM with a custom-designed reasoning module.

The Coconut framework is designed to enhance the reasoning capabilities of LLMs. It consists of a pre-trained LLM and a custom-designed reasoning module. The reasoning module is trained using a novel architecture that combines a pre-trained LLM with a custom-designed reasoning module. This module enables the model to generate coherent and logical reasoning paths for complex tasks, significantly improving its performance on benchmarks that require deep understanding and logical deduction. The proposed method is evaluated on various tasks, demonstrating superior results compared to existing baselines.

The Coconut framework is designed to enhance the reasoning capabilities of LLMs. It consists of a pre-trained LLM and a custom-designed reasoning module. The reasoning module is trained using a novel architecture that combines a pre-trained LLM with a custom-designed reasoning module. This module enables the model to generate coherent and logical reasoning paths for complex tasks, significantly improving its performance on benchmarks that require deep understanding and logical deduction. The proposed method is evaluated on various tasks, demonstrating superior results compared to existing baselines.

The Coconut framework is designed to enhance the reasoning capabilities of LLMs. It consists of a pre-trained LLM and a custom-designed reasoning module. The reasoning module is trained using a novel architecture that combines a pre-trained LLM with a custom-designed reasoning module. This module enables the model to generate coherent and logical reasoning paths for complex tasks, significantly improving its performance on benchmarks that require deep understanding and logical deduction. The proposed method is evaluated on various tasks, demonstrating superior results compared to existing baselines.

The Coconut framework is designed to enhance the reasoning capabilities of LLMs. It consists of a pre-trained LLM and a custom-designed reasoning module. The reasoning module is trained using a novel architecture that combines a pre-trained LLM with a custom-designed reasoning module. This module enables the model to generate coherent and logical reasoning paths for complex tasks, significantly improving its performance on benchmarks that require deep understanding and logical deduction. The proposed method is evaluated on various tasks, demonstrating superior results compared to existing baselines.

The Coconut framework is designed to enhance the reasoning capabilities of LLMs. It consists of a pre-trained LLM and a custom-designed reasoning module. The reasoning module is trained using a novel architecture that combines a pre-trained LLM with a custom-designed reasoning module. This module enables the model to generate coherent and logical reasoning paths for complex tasks, significantly improving its performance on benchmarks that require deep understanding and logical deduction. The proposed method is evaluated on various tasks, demonstrating superior results compared to existing baselines.

The Coconut framework is designed to enhance the reasoning capabilities of LLMs. It consists of a pre-trained LLM and a custom-designed reasoning module. The reasoning module is trained using a novel architecture that combines a pre-trained LLM with a custom-designed reasoning module. This module enables the model to generate coherent and logical reasoning paths for complex tasks, significantly improving its performance on benchmarks that require deep understanding and logical deduction. The proposed method is evaluated on various tasks, demonstrating superior results compared to existing baselines.

The Coconut framework is designed to enhance the reasoning capabilities of LLMs. It consists of a pre-trained LLM and a custom-designed reasoning module. The reasoning module is trained using a novel architecture that combines a pre-trained LLM with a custom-designed reasoning module. This module enables the model to generate coherent and logical reasoning paths for complex tasks, significantly improving its performance on benchmarks that require deep understanding and logical deduction. The proposed method is evaluated on various tasks, demonstrating superior results compared to existing baselines.

- æ??ç« i¼?Texti¼?
â??ã??ã??ã??¼ã?ã?³â??
- é?çæ?£ç??ã??ã??¼ã?ã?³â??i¼?ã??Aã?•ã??ã?•ã?•ã??Bã?•ã??ã?•ã?•...i¼?
â??ã??ã??æ?•
- â?|ç?•â??é?èj"ç?¾i¼?Hidden state 0i¼?
â??
- Hidden State1 + Hidden State2 + Hidden State 3ã??i¼?è"èª?â??ã?•ã?ªã?
?æ½?â?"ç??ã?ªæ?•è??ã??ã, |ã?¹ã??i¼?
â??
- æ??çµ?â?°â??

ã?ã?ã?â½çã?«ã?ã??ã?¾ã?ã??

äºé??ã?®éã?®ä,ã?§ã?æ??ã?«æ?³èµ.ã?ã??ã??æ?•è??ã?®âã?•4ã?®é?â??ã??ã?•ã?ã?§æ?•è??ã?
«â¹ã?â?°ã??ã?ã?ã?«ã?ã??ã?¾ã?ã??

ã?ã?¥ã?¼ã??ã?³ã?°ã?®é??ã?«ã?ã?ã?é?ç?¶æ?ã??ã?«ã?ã?ã??ã?¼ã?¿ã?»ã??ã??ã??ç?æ?ã?
ã??ã?ã?ã?«ã?ã??ã?æ?³ã?ã?§ã?ã?¾ã?ã??

äºé??ã?®æ?•è??é?ç?ã?ã?®â¾æ?

äºé??ã?®æ?•è??é?ç?ã?ã?®â¾â¿é?çä¿ã?ã?ã??ã?æ??ã?ã?«ã?ã??ã?§ã?ã??ã?ã??

äºé??ã?®æ?
è??æ®µé??

LLMâ? é?"ã?®â¾â¿èç´

èª-æ??

â? ç?

¾è±jã??è"æ?¶ã

•®æ?ç??ã?éã? hidden state

«â??ã? ç¾i¼?æ½?â?ã??ã?ã??ã?«i¼?

¼ã??ã??ã??â?•

æµ®ã?ã?¶

â?jã?ã?ã?ªã?ã?

®â??è£?ã??ã??ã

•ã??ã?ã?ªã?ã? latent reasoning loopi¼?é?£é??ã?

ã??ã?ã?ªã?ã? ?ã?? hidden statei¼?

"æ"è¼?ã?ã??

â?çæ?¹â?æ?§ã?

?â?°ã?¾ã?£ã?jã?ã? latent â?? token decoding

•ã??ã?è"è??ã?

§ã?¾ã?ã?ã??

â?£ã?»è?ã?â

±æ??ã?ã??ã?

?ã?ã?®èj"ç?¾ã? natural language generation

«æ?´ã?ã??

è"è??ã?«ã?ã??â?ã?

®é?£æ?³ç??èj"è±jã??ã?ªã?jã?¼ã?,ã?»æ??

ã?»æ§é?ã?®æ?â?ç?¶æ?ã??

èª?æ?°ã?®â??è£?æ?¹â?ã??â?æ??ã?«ä¿æ

•ã?â? ç??ã?«æ?çç´çã??èj?ã?ã??

æ½?â?"ç??æ?•è??ã?è"èª?ã??ã?ã??ã?

âª?â??ã?ã??ã??ã??

â¾è©±ã?»æ??æ?,ã?»èª-æ??ã?ã?jã?

®æ??çµ?â?°â??ã??

Coconutã?ã?ã?ã?®ã??â?ã?â?jã?ã?ã?®é?â??â??â??

ã?ª¾ã??ã?ºé??ã?§è"ã?ã?â??ã?¾ã?è"è??ã?«ã?ã?£ã?jã?ã?ã?æ?•è??æ®µé??â?â??â??ã??

Alã?®â?é?"ã?§ã?ã??ã?¥ã?-ã?¼ã??ã?ã??ã?ã?ã?ã?jã?ã??ã?ã?ã?§ã?ã??

å¼æ¥ã•®LLMã¬ã?â?¢ã¬â?£ã•®ã??è?è?â?â?®æ®é??ã?ã?ã?æ?±ã?ã?ã?ã?£ã?ã??
ã•ã?ã??ã?â??ã?ã?ã?«ã?ã??ã?ã??å§ã?ã?ã?ã?ã?Alâ?ã•ã?£ã?ã??

ã?ã?ã?ã?ä°é??ã¬ã?

ã??ã¼ã¬è??è??ã«ã?
ã¬ã?ã?ã?æ??é??ã??æ?ã?£ã¬ã?ã??ã??

Coconutã¬ã?ã¼ã?
ã??ã?ã??ã?«ã®é?£ç¶?ã?

KAGURAã¬ã?ã?ã?ã?æ¬?

ZIKUUã¬æ?ã?
KAGURAã¬ã?

KAGURAã¬é??ã?ç?¥ã??æ?±ã?
¬Coconutã¬ã¼ã¼ã¬ã?ã?ã?ã?ã?ã?
Coconutã¬ã??ã?é?è?ç?ã?
ã?
ã?
ã?

Coconutã¬ã?

è?ç?¹	Coconut	KAGURA
æ?è??ã• ®ä½ç½®	ã?¢ã??ã?«ã? é?¬ã• ®æ½?ã?¬ç®é?? ã? é?¬ã??ã?¬ã??ã?«ã	ã? ®é??ã?¬½?
ä,ã??æ?§ã• ®ã®?ç¼	•ã?æ??ã• «è?æ?æ?¹ã• ã??ã?ã?	è?æ?ã?ã?ã?ã?ã?ã?ã?ã?ã?ã?ã?ã? ?A/B/Ci¼ã??ã?ã?ã?ã?ã?ã?ã?
æ¬jã•®æ?• è??	hidden stateã??ã?ã? ¥ã??ã?ã?ã?ã?ã?	ç??æ?çµæ??ã??æ¬è¼ã?»ã?çµ±ã?ã?ã?ã?ã?ã?ã?
æ¬è³ç??ã •ç??ã?	è?è?ã?ã?ã?ã?ã?ã? ®æµã??ç??æ?¬è«? ä,ã?ã?ã? ®ç?ç?ç??æ?•	è?è?ã?ã?ã?ã?ã?ã?ã?ã?ã?ã?ã?ã? è?ç?¹ç??ã?ã?ã?ã?ã?
ã ±é??ç?¹	è??ã??è¶ã?ã? ã?æ?¹ã?æ?§ã??ã? ã??æ?»ã?	

Coconutã¬æ²é?»ã®ä,ã?§è??ã?
jã??ã??ã?ç?ç??ã?æ££æ?ç?¢ã??ã??ã?ã?ã?ã?ã?ã?ã?ã?ã?ã?ã?ã?ã?ã?ã?ã?
§ã¼ã¼ã¬ã?ã?ã?ã?ã?

[illegible][illegible]

ã•?ã??ã-ã??é è?³ã??äºµæ•?ã•?ã|ç?ºã±?ã•§ã•ã???ã ±ã•?ä½?ã?ã•®ã??ã•?ã•ªä»?çµ?ã•¿ã•ã??è"?ã•
?ã??ã•?ã??ã•?ã??ã•¾ã•?ã??ã??

[illegible][illegible]

å®@åµ§å·¥ã·?â°çµ¼ã??å»°ç-?ã?ã??é??ã«ç?·ã·?ã??æ??è;?ã·-ã·è·ã·£ã·?æ?±ã·ã·?ã??ä°µæ·?ã·
§ã·ã·?æ§?é? ä??ä½?ã??æ??è;?ã·§ã·?ã??ã·ã??ã··ã·?ã·?ç?°æ?³ã·§ã·?ã??

ä¼•çµ±ã•æ?æ?ã?ã??ã•ã•ã?ã??ç¶?æ?¿ã•ã??ã??ã??
ç¶?æ?¿ã•ã??ã??ã?»?çµ?ã•¿ã??ä½?ã?ã•ã??

ã?¢ã??ã?¥ã?ã??ã?«ã??æ?³ã?¿ è|ã??
ã?ã??ã?ZIKUUæµã?§ã?ã??

Category

1. æ?¥ã? ã•®æ'»å??

Tags

1. AI
2. Coconut
3. KAGURA
4. LLM
5. ä ?é??å??å ?

6. å±±æ¢“ç??

Date Created
2025å¹´11æ??7æ?¥
Author
kazuo-tsubaki

default watermark