LLaDA: Large Language Diffusion Models
Autoregressive   LLM

Ken

- ：LLM

- : LLM
- : LLaDA masking diffusion 8B LLaMA3 8B

- : LLM
- : LLaDA   masking diffusion 8B   LLaMA3 8B
- :   in-context learning instruction-following

- AR LLM

-
- AR LLM
- LLaDA   Eq.3, Eq.4   remask

- ▶
- ▶   AR   LLM
- ▶ LLaDA   Eq.3, Eq.4   remask
- ▶

- :

$$p(x) = \prod_{i=1}^{L} p(x_i \mid x_{<i})$$

- ▶ :
- ▶ : AR

$$p(x) = \prod_{i=1}^{L} p(x_i \mid x_{<i})$$

- ▶ :
- ▶ :   AR
- ▶   : 8B

$$p(x) = \prod_{i=1}^{L} p(x_i \mid x_{<i})$$

- LM　　　　GPT-2

- LM   GPT-2
-

- ▶ LM　　　　　　　GPT-2
- ▶
- ▶ masked　　: MaskGIT

LLaDA

▶ Forward: t $t \sim U[0, 1]$

LLaDA

▶ Forward: t $t \sim U[0,1]$
▶ Reverse:

LLaDA

- Forward: t $t \sim U[0,1]$
- Reverse:
- : Causal mask   Transformer

LLaDA

- Forward: t $t \sim U[0,1]$
- Reverse:
-    : Causal mask   Transformer

- AR   1

LLaDA

- Forward: t $t \sim U[0, 1]$
- Reverse:
- : Causal mask Transformer

- AR 1
- LLaDA

LLaDA

▶ Forward:　　　t　　$t \sim U[0,1]$
▶ Reverse:
▶ 　: Causal mask　Transformer

▶ AR　　1
▶ LLaDA
▶

masked token CE :

$$\mathcal{L}(\theta) = -\mathbb{E}_{t,x_0,x_t} \left[\frac{1}{t} \sum_{i=1}^{L} \mathbf{1}[x_t^i = M] \log p_\theta(x_0^i \mid x_t) \right]$$

:

$$-\mathbb{E}_{p_{\text{data}}(x_0)}[\log p_\theta(x_0)] \leq \mathcal{L}(\theta)$$

▶

masked token CE :

$$\mathcal{L}(\theta) = -\mathbb{E}_{t,x_0,x_t} \left[\frac{1}{t} \sum_{i=1}^{L} \mathbf{1}[x_t^i = M] \log p_\theta(x_0^i \mid x_t) \right]$$

:

$$-\mathbb{E}_{p_{\mathsf{data}}(x_0)}[\log p_\theta(x_0)] \leq \mathcal{L}(\theta)$$

▶
▶

- Pre-train: 2.3T tokens 1B/8B   seq length 4096 Warmup-Stable-Decay

- ▶ Pre-train: 2.3T tokens 1B/8B   seq length 4096 Warmup-Stable-Decay
- ▶ SFT: 4.5M prompt-response pairs prompt response

- Pre-train: 2.3T tokens 1B/8B   seq length 4096 Warmup-Stable-Decay
- SFT: 4.5M prompt-response pairs prompt   response
- Sampling: step

- Pre-train: 2.3T tokens 1B/8B   seq length 4096 Warmup-Stable-Decay
- SFT: 4.5M prompt-response pairs prompt   response
- Sampling: step
- Sampling   remask low-confidence remasking

- : ARM baseline LLaDA

- : ARM baseline LLaDA
- : FLOPs 6

- ▶ : ARM baseline LLaDA
- ▶ : FLOPs 6
- ▶ : ARM

- ▶ : ARM baseline LLaDA
- ▶ : FLOPs 6
- ▶ : ARM
- ▶ : MMLU/GSM8K LLaDA

- ▶ : ARM baseline LLaDA
- ▶ : FLOPs 6
- ▶ : ARM
- ▶ : MMLU/GSM8K LLaDA
- ▶ : AR

8B 模型 Base 版本性能对比

- LLaDA 8B Base vs LLaMA2 7B Base

- LLaDA 8B Base vs LLaMA2 7B Base
- 15

8B 模型 Base

- LLaDA 8B Base vs LLaMA2 7B Base
- 15 个任务
- LLaDA 8B Base vs LLaMA3 8B Base

- LLaDA 8B Base vs LLaMA2 7B Base
- 15
- LLaDA 8B Base vs LLaMA3 8B Base
-

- LLaDA 8B Base vs LLaMA2 7B Base
- 15
- LLaDA 8B Base vs LLaMA3 8B Base
-
- Table 1

- LLaDA 8B Base vs LLaMA2 7B Base
- 15
- LLaDA 8B Base vs LLaMA3 8B Base
- ▶
-     Table 1
- GSM8K: 70.3 (LLaDA) vs 48.7 (LLaMA3 8B Base)

- LLaDA 8B Base vs LLaMA2 7B Base
- 15
- LLaDA 8B Base vs LLaMA3 8B Base
-
- 　 Table 1
- GSM8K: 70.3 (LLaDA) vs 48.7 (LLaMA3 8B Base)
- MATH: 31.4 vs 16.0

- LLaDA 8B Base vs LLaMA2 7B Base
- 15
- LLaDA 8B Base vs LLaMA3 8B Base
-
- Table 1
- GSM8K: 70.3 (LLaDA) vs 48.7 (LLaMA3 8B Base)
- MATH: 31.4 vs 16.0
- HumanEval: 35.4 vs 34.8

SFT

- LLaDA SFT RL instruction-following

SFT

- LLaDA SFT RL instruction-following
-

SFT

- LLaDA SFT RL instruction-following
-
- LLaMA3 Instruct

SFT

- LLaDA SFT RL instruction-following
-
-     LLaMA3 Instruct
- : AR

Reversal Curse

- :   forward   reversal

Reversal Curse

- :   forward   reversal
- Table 4

Reversal Curse

- : forward reversal
- Table 4
- GPT-4o: Forward 82.7 / Reversal 34.3

Reversal Curse

- : forward reversal
- Table 4
- GPT-4o: Forward 82.7 / Reversal 34.3
- Qwen2.5-7B: 75.9 / 38.0

Reversal Curse

- : forward reversal
- Table 4
- GPT-4o: Forward 82.7 / Reversal 34.3
- Qwen2.5-7B: 75.9 / 38.0
- LLaDA-8B Instruct: 51.8 / 45.6

Reversal Curse

- ▶ : forward reversal
- ▶ Table 4
- ▶ GPT-4o: Forward 82.7 / Reversal 34.3
- ▶ Qwen2.5-7B: 75.9 / 38.0
- ▶ LLaDA-8B Instruct: 51.8 / 45.6
- ▶ : LLaDA forward reversal

Reversal Curse

- ▶ :   forward   reversal
- ▶   Table 4
- ▶ GPT-4o: Forward 82.7 / Reversal 34.3
- ▶ Qwen2.5-7B: 75.9 / 38.0
- ▶ LLaDA-8B Instruct: 51.8 / 45.6
- ▶   : LLaDA forward   reversal
- ▶   :

3

- 8B　　LM　　LLM

3

▶ 8B　　LM　　　LLM
▶

3

- 8B LM LLM
-
- reversal reasoning AR

- $< 10^{23}$ FLOPs

- $< 10^{23}$ FLOPs
-

- $< 10^{23}$ FLOPs
-
- RL

- $< 10^{23}$ FLOPs
-
- RL
- KV cache

- $< 10^{23}$ FLOPs
-
- RL
- KV cache
-

▶ : AR

- : AR
- : remask step CFG

- : AR
- : remask step CFG
- : RLHF/RLAIF agent AR

- LLaDA　AR LLM　　8B

- LLaDA AR LLM 8B
- reversal

- LLaDA AR LLM 8B
- reversal
- RL AR

- LLaDA AR LLM 8B
- reversal
- RL AR
- : AR

Q&A

Q1. BERT
A. mask $t \sim U[0,1]$

Q2. reversal
A.

Q3.
A. 1 1step AR step - sampling /

- Paper: Large Language Diffusion Models (Nie et al.)

- Paper: Large Language Diffusion Models (Nie et al.)
- Project: https://ml-gsai.github.io/LLaDA-demo/

- Paper: Large Language Diffusion Models (Nie et al.)
- Project: https://ml-gsai.github.io/LLaDA-demo/
- Local source PDF:
  ../_papers/LargeLanguageDiffusionModels.pdf