Metric supervision enables LLMs to learn structured **token geometry**.

which is useful for…

Reward Modelling

Input: [Instruction]
[Sample response]
Output: Helpfulness: 4,
Safety:1,
Verbosity: 3,...

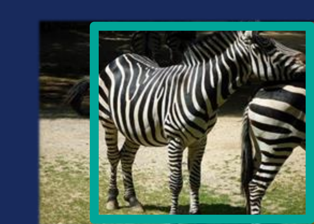
Robotic Manipulation

Input: Put into the

Output: Move arm to (0.2,0.4)
and Rotate 108°

Visual Grounding

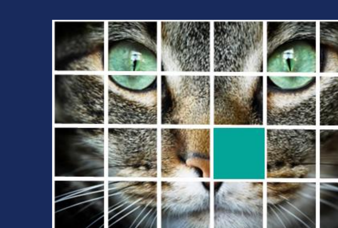
Input: Detect the full zebra.



Output: The full zebra is at
[(49,10),(540,465)]

Text-to-Image

Input: Draw a cat.



Output: <IMG28>, <IMG10>
<IMG32>.....

Motivation

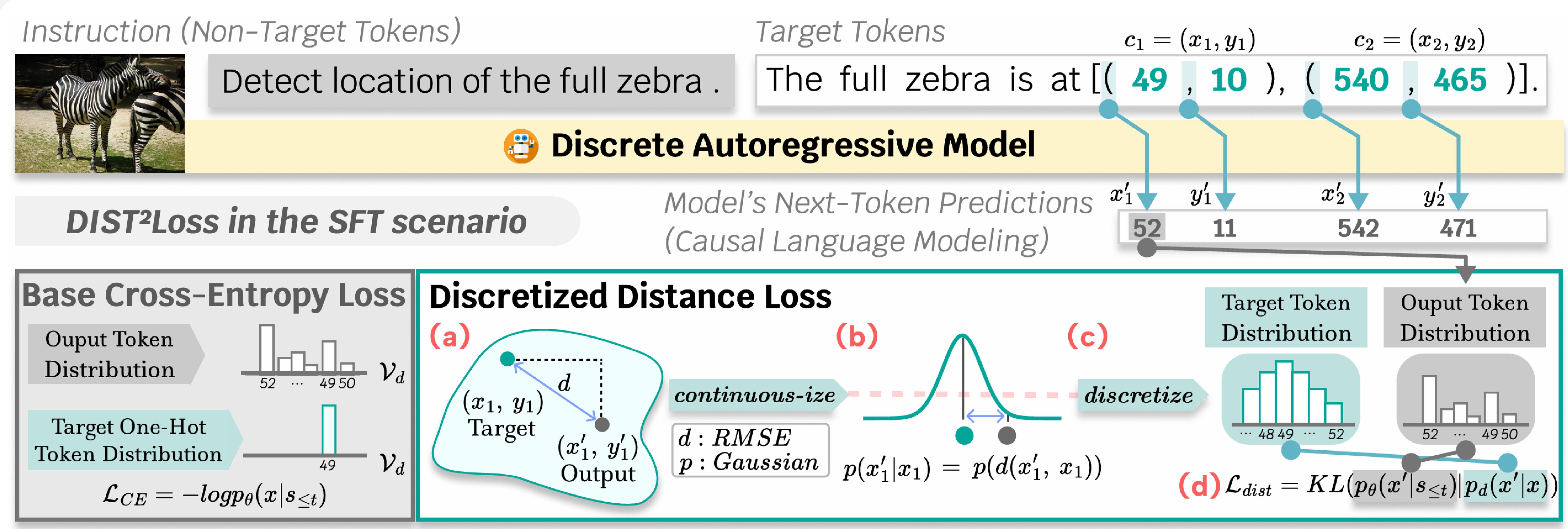
- Standard autoregressive models do not learn that nearby tokens can be better than distant ones.
- Our goal is to teach LLMs the geometry of structured tokens, such as coordinates, angles.

Methods

- We augment one-hot CE with **a distance-aware soft target distribution**.

$$p_d(v \mid x, t) = \frac{\exp(-d(v, x, t)/\tau)}{\sum_{v' \in \mathcal{V}_d} \exp(-d(v', x, t)/\tau)}$$

- Tokens closer to the target under metric $d(\cdot,\cdot)$ receive higher probability mass.



Results

Reward Modelling

Models	RewardBench Chat	Chat Hard	Safety	Reasoning	Average
Tulu-v2.5-RM-13B	93.6	68.2	77.3	88.5	81.9
GPT-3.5	92.2	44.5	65.5	59.1	65.3
Claude-3-haiku	73.7	92.7	52.0	79.5	70.6
Prometheus-2-7B	85.5	49.1	77.1	76.5	72.0
Llama-*binary*	83.8	34.7	39.9	73.5	58.0
Llama-*sft*	89.1	49.3	79.2	83.9	75.3
Llama-*dist*	95.0 (↑**4.9**)	69.1 (↑**19.8**)	86.5 (↑**7.3**)	90.4 (↑**6.5**)	85.3 (↑**10.0**)

Robotic Manipulation

	L1			L2		
#Data	1K	10K	100K	1K	10K	100K
RT-2	1.9	21.9	73.1	3.8	17.7	70.4
LLaRA-*sft*	49.6	82.3	88.5	46.2	78.1	84.6
LLaRA-*vocab*	50.8	81.0	87.0	44.6	77.2	83.5
LLaRA-*dist*	**53.9**	**83.4**	**89.5**	**51.5**	**82.8**	**86.1**

Visual Grounding

Models	#PT	#FT	RefCOCO val	test-A	test-B
Ferretv2	1.1M	127K	92.8	94.7	88.7
Florence-2-B	126M	127K	92.6	94.8	91.5
Florence-2-L	126M	127K	93.4	95.3	92.0
Phi3V-*sft*	0	127K	94.3	93.5	86.0
Phi3V-*vocab*	0	127K	94.5 (↑0.2)	93.2 (↓0.3)	86.0 (–)
Phi3V-*dist*	0	127K	94.8 (↑**0.5**)	94.5 (↑**1.0**)	87.3 (↑**1.3**)

Text-to-Image

Models	Epoch #Params	50 FID ↓	50 IS ↑	Full (300) FID ↓	Full (300) IS ↑
VQGAN	227M	-	-	18.65	80.4
VQGAN	1.4B	-	-	15.78	74.3
LlamaGen-*sft*	111M	10.03	116.37	6.44	157.17
LlamaGen-*dist*	111M	9.41	127.44	6.27	164.32
LlamaGen-*sft*	343M	4.24	206.74	3.08	256.07
LlamaGen-*dist*	343M	**4.18**	**209.41**	**3.04**	**258.19**

Across various downstream tasks, Adding Discretized distance loss improves performance, with particularly strong gains in **low-data regimes**

Teaching Metric Distance to Discrete Autoregressive Language Models

Jiwan Chung, Saejin Kim, Yongrae Jo, Jaewoo Park, Dongjun Min, and Youngjae Yu

Jiwan.chung.research@gmail.com

Metric supervision enables LLMs to learn structured **token geometry**.

which is useful for...

Reward Modelling

Input: [Instruction]
[Sample response]

Output: Helpfulness: 4,
Safety:1,
Verbosity: 3,...

Robotic Manipulation

Input: Put into the

Output: Move arm to (0.2,0.4)
and Rotate 108°

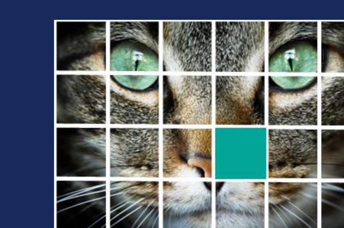
Visual Grounding

Input: Detect the full zebra.

Output: The full zebra is at
[(49,10),(540,465)]

Text-to-Image

Input: Draw a cat.



Output: <IMG28>, <IMG10>
<IMG32>.....

LG AI Research

Motivation

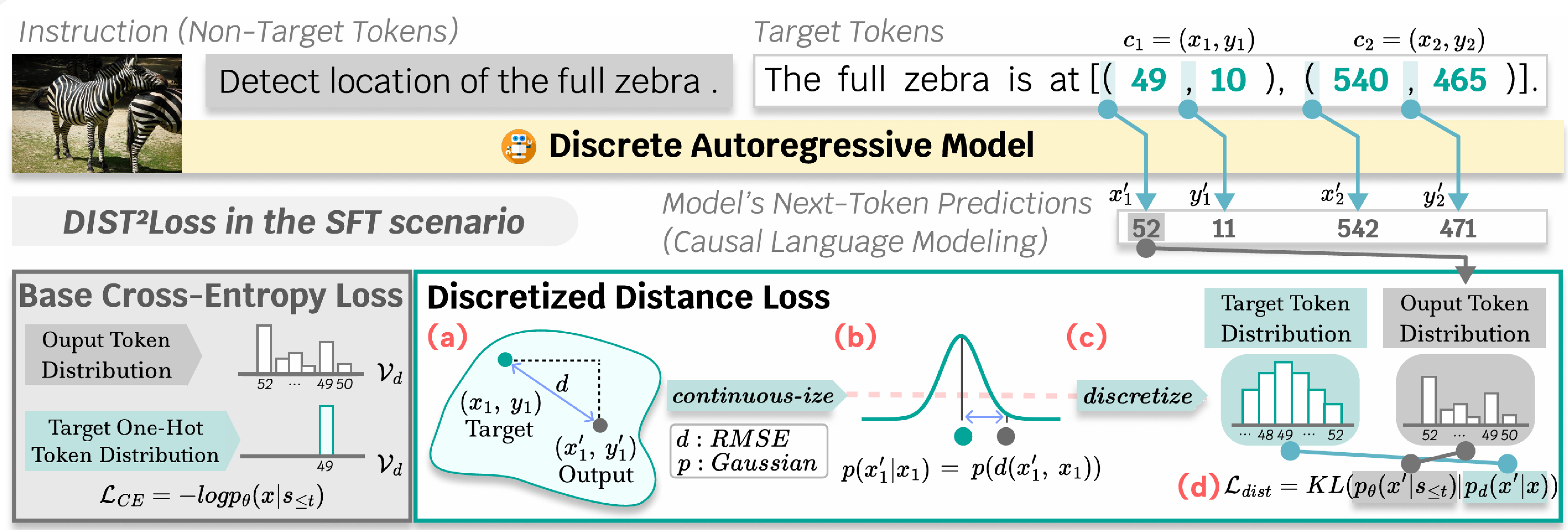
- Standard autoregressive models do not learn that nearby tokens can be better than distant ones.
- Our goal is to teach LLMs the geometry of structured tokens, such as coordinates, angles.

Methods

- We augment one-hot CE with **a distance-aware soft target distribution**.

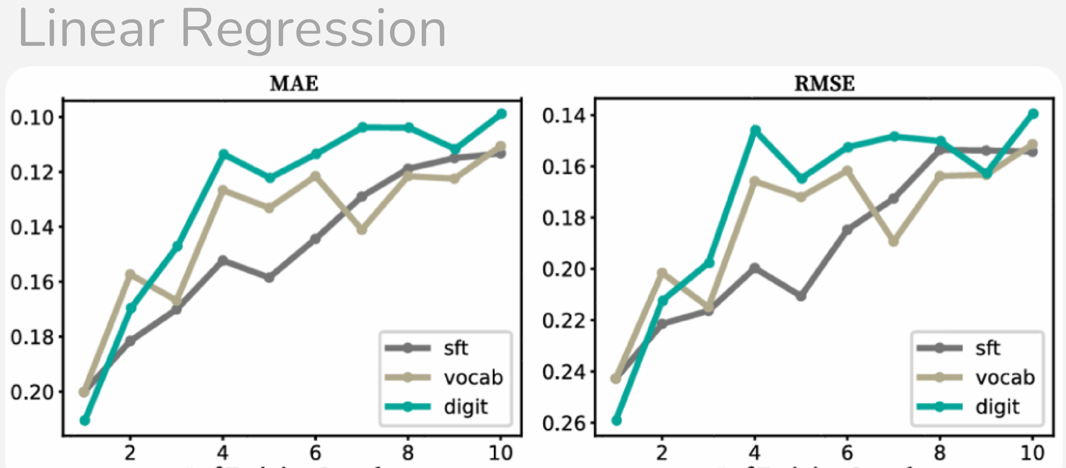
$$p_d(v \mid x, t) = \frac{\exp(-d(v, x, t)/\tau)}{\sum_{v' \in V_d} \exp(-d(v', x, t)/\tau)}$$

- Tokens closer to the target under metric $d(\cdot,\cdot)$ receive higher probability mass.



Results

Linear Regression



Reward Modelling

Models	RewardBench Chat	Chat Hard	Safety	Reasoning	Average
Tulu-v2.5-RM-13B	93.6	68.2	77.3	88.5	81.9
GPT-3.5	92.2	44.5	65.5	59.1	65.3
Claude-3-haiku	73.7	92.7	52.0	79.5	70.6
Prometheus-2-7B	85.5	49.1	77.1	76.5	72.0
Llama-*binary*	83.8	34.7	39.9	73.5	58.0
Llama-*sft*	89.1	49.3	79.2	83.9	75.3
Llama-*dist*	95.0 (↑**4.9**)	69.1 (↑**19.8**)	86.5 (↑**7.3**)	90.4 (↑**6.5**)	85.3 (↑**10.0**)

Robotic Manipulation

#Data	L1 1K	L1 10K	L1 100K	L2 1K	L2 10K	L2 100K
RT-2	1.9	21.9	73.1	3.8	17.7	70.4
LLaRA-*sft*	49.6	82.3	88.5	46.2	78.1	84.6
LLaRA-*vocab*	50.8	81.0	87.0	44.6	77.2	83.5
LLaRA-*dist*	**53.9**	**83.4**	**89.5**	**51.5**	**82.8**	**86.1**

Visual Grounding

Models	#PT	#FT	RefCOCO val	test-A	test-B
Ferretv2	1.1M	127K	92.8	94.7	88.7
Florence-2-B	126M	127K	92.6	94.8	91.5
Florence-2-L	126M	127K	93.4	95.3	92.0
Phi3V-*sft*	0	127K	94.3	93.5	86.0
Phi3V-*vocab*	0	127K	94.5 (↑0.2)	93.2 (↓0.3)	86.0 (–)
Phi3V-*dist*	0	127K	94.8 (↑**0.5**)	94.5 (↑**1.0**)	87.3 (↑**1.3**)

Across various downstream tasks, Adding Discretized distance loss improves performance, with particularly strong gains in **low-data regimes**

Teaching Metric Distance to Discrete Autoregressive Language Models

Jiwan Chung, Saejin Kim, Yongrae Jo, Jaewoo Park, Dongjun Min, and Youngjae Yu

Jiwan.chung.research@gmail.com