



Disambiguation of Idiomatic Expressions with Images

Katsumi Ibaraki

Motivation

▶ Emergence of CLIP

- ▶ Contrastive Language-Image Pre-training
- ▶ Neural network that learns visual concepts from natural language supervision
- ▶ Originally for visual classification

Motivation

▶ Uses of CLIP

- ▶ DALL-E and Stable Diffusion
 - ▶ Generate digital images from natural language descriptions

▶ Issues

- ▶ Disambiguation of entities
 - ▶ e.g. financial bank vs river banks



DALL-E 2: An astronaut riding a horse in photorealistic style.

“Break a leg”



Task

- ▶ **Explore approaches to generate images that expresses the figurative meaning**
 - ▶ Compare natural language-based approaches with text-image-based approaches
- ▶ **Generate images for 10 English idioms**

Methodology

- ▶ **Baseline and Fine-tuned model**
 - ▶ Stable Diffusion model
 - ▶ Stable Diffusion model fine-tuned on other idioms
- ▶ **Rephrasing the expressions**
 - ▶ GPT-3.5
 - ▶ Convert idiom into simpler expressions

Experiments

- ▶ **Generations**

- ▶ Each model given the 10 idioms to generate 20 images

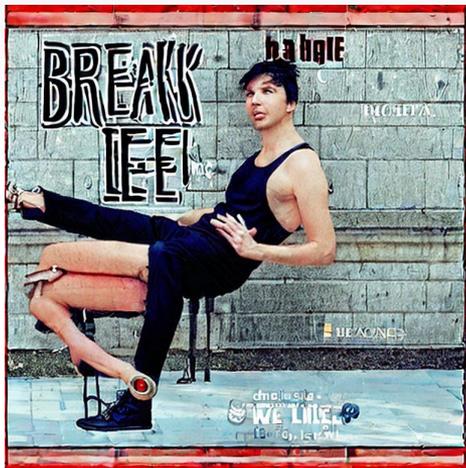
- ▶ **Evaluation**

- ▶ 2 independent evaluators asked to judge:
 - ▶ Likely literal or likely figurative

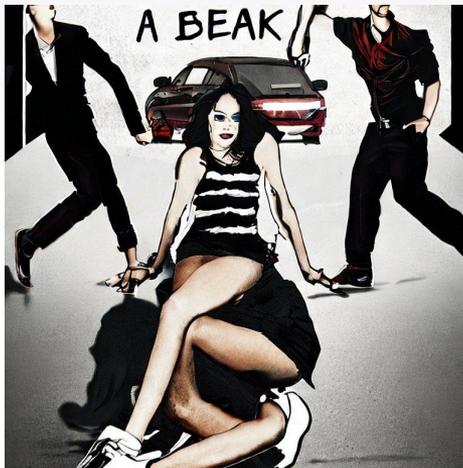
Results

Idiomatic Expression	Untuned		Finetuned		GPT-Rephrased	
	Literal%	Figurative%	Literal%	Figurative%	Literal%	Figurative%
Bite the bullet	1.00	0.00	0.90	0.10	0.15	0.85
Break a leg	0.95	0.05	0.80	0.20	0.10	0.90
Break the ice	1.00	0.00	1.00	0.00	0.00	1.00
Get out of hand	0.95	0.05	0.90	0.10	0.05	0.95
Hit the sack	1.00	0.00	0.85	0.15	0.00	1.00
On the ball	0.80	0.20	0.70	0.30	0.10	0.90
Pull yourself together	0.95	0.05	0.85	0.15	0.20	0.80
To get bent out of shape	0.90	0.10	0.85	0.15	0.10	0.90
Under the weather	1.00	0.00	0.75	0.25	0.00	1.00
Wrap your head around something	1.00	0.00	0.90	0.10	0.00	1.00
Average	0.955	0.045	0.850	0.150	0.070	0.930

“Break a leg”



(a) Untuned



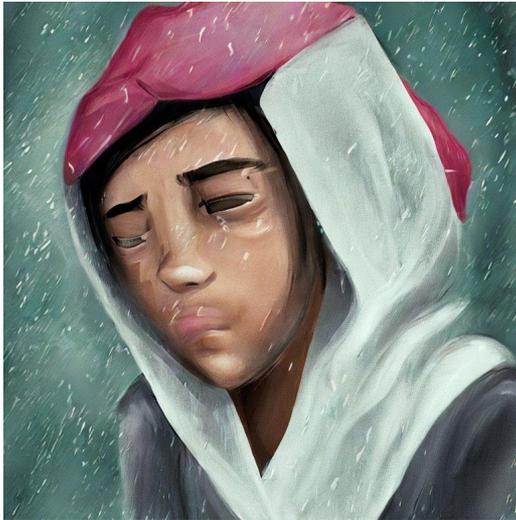
(b) Fine-tuned



(c) Rephrased

Discussion

- ▶ **Effects of fine-tuning**
 - ▶ “Under the weather” (sick)



Discussion

▶ **Future work**

- ▶ Combine rephrasing with fine-tuning
- ▶ Compare DALL-E and Stable Diffusion
- ▶ Extend past idioms
 - ▶ Provide two similar expressions and differentiate by images