Proposal: Exploring Visual Metaphors and Sentiments: A Multimodal Approach to Meme Captioning

Dataset Preparation
We will use the MEMECAP dataset from the provided GitHub repository (https://github.com/eujhwang/meme-cap), preprocess the data as needed, such as resizing images, tokenizing captions, and extracting visual metaphors.

Model Selection
We can choose appropriate vision and language models for the task, like Open Flamingo, MiniGPT4, and LLaMA.

Model Training
Fine-tune the selected models on the MemeCap dataset. Experiment with different training setups, such as zero-shot, few-shot, and fine-tuning, as mentioned in the paper.

Caption Generation
Develop a system that takes a meme image and corresponding title (representing what the user is trying to convey) as input and generates a caption that captures the visual metaphor and the intended meaning.
We are also seeking new approaches that might help improve performance. For example, retrieval-augmented language models (RALMs) have shown impressive performance on knowledge-intensive tasks. Simply prepending retrieved documents to the input without updating the language models (LMs) could help. It is possible that some discussion among the Reddit threads reveals useful information that can help the model interpret the memes better.

Possible Evaluation Metrics
Evaluate the generated captions using metrics like BLEU, ROUGE, and BERTScore, as used in the MemeCap paper (https://arxiv.org/abs/2305.13703).
Consider conducting human evaluations to assess the correctness, appropriateness, and completeness of the captions.

Iterate and Improve
Based on the evaluation results, iterate on your model and training process to improve performance.
Explore different model architectures, training strategies, and input representations to enhance the captioning quality.

Some applications of this project
1.We could analyze the generated captions to extract insights related to public sentiment and perceptions. 2.With precise image captioning, it is possible to suggest proper memes for discussion or even generate new memes based on different cultures.

Youjun Chen, yc7093@nyu.edu
Yanfeng Du, yd2727@nyu.edu
Kexin Xu, kx2139@nyu.edu
PangLi Yang, py2236@nyu.edu