

# Persuasive Techniques

Dataset

Meme Image



Text Written in  
the Meme

LLaVA-1.5-7B

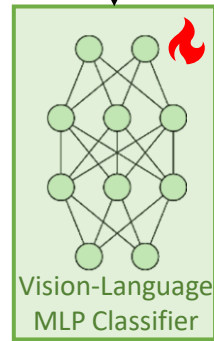
Caption for  
the Meme



BERT Large



Vision Encoder  
(CLIP-ViT-L)



Vision-Language  
MLP Classifier

Classification  
Label(s)



## Meme Image

Me leaving Area 51 with a shopping trolley that has four working wheels:



Text Written in  
the Meme

GPT4-V(ision)



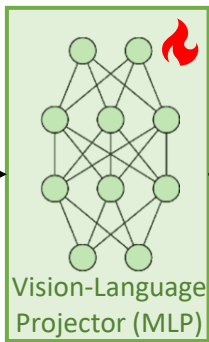
Classification  
Label(s)

## Meme Image

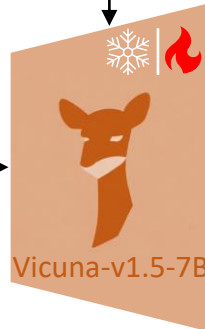
Me leaving Area 51 with a shopping trolley that has four working wheels:



Text Written in  
the Meme



LLaVA-1.5-7B



Classification  
Label(s)

## Meme Image

Me leaving Area 51 with a shopping trolley that has four working wheels:



Text Written in  
the Meme

GPT4-V(ision)



Caption for  
the Meme



Vicuna-v1.5-7B

Classification  
Label(s)

## Meme Image

Me leaving Area 51 with a shopping trolley that has four working wheels:

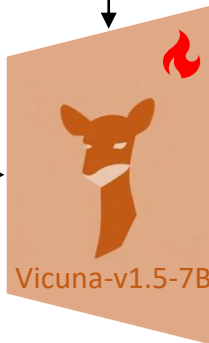


Text Written in  
the Meme

GPT4-V(ision)



Caption for  
the Meme



Classification  
Label(s)

## Meme Image

Me leaving Area 51 with a shopping trolley that has four working wheels:



Text Written in  
the Meme

LLaVA-1.5-7B

Caption for  
the Meme



Visual BERT

Classification  
Label(s)

## Meme Image

Me leaving Area 51 with a shopping trolley that has four working wheels:



GPT4-V(ision)



Caption for  
the Meme

Text Written in  
the Meme



BERT/roBERTa  
Large

Classification  
Label(s)

## Meme Image

Me leaving Area 51 with a shopping trolley that has four working wheels:



Text Written in  
the Meme

LLaVA-1.5-7B

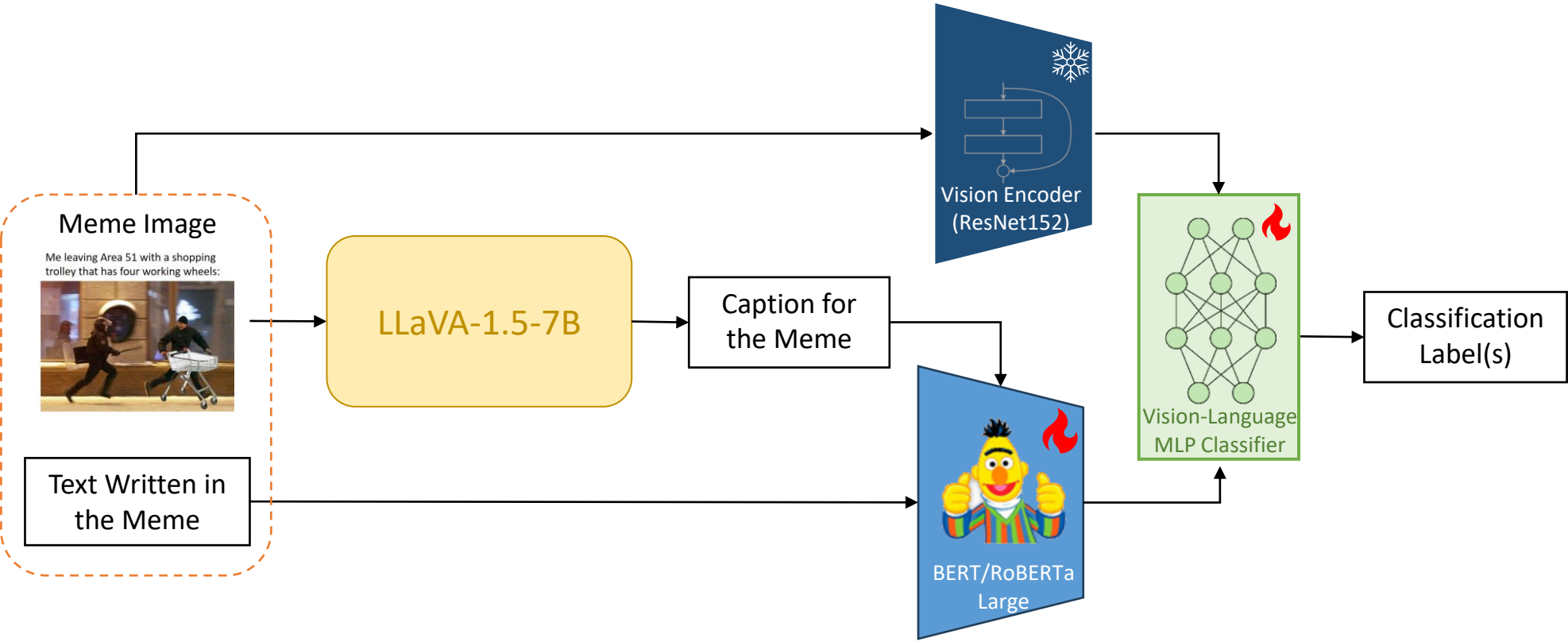
Caption for  
the Meme

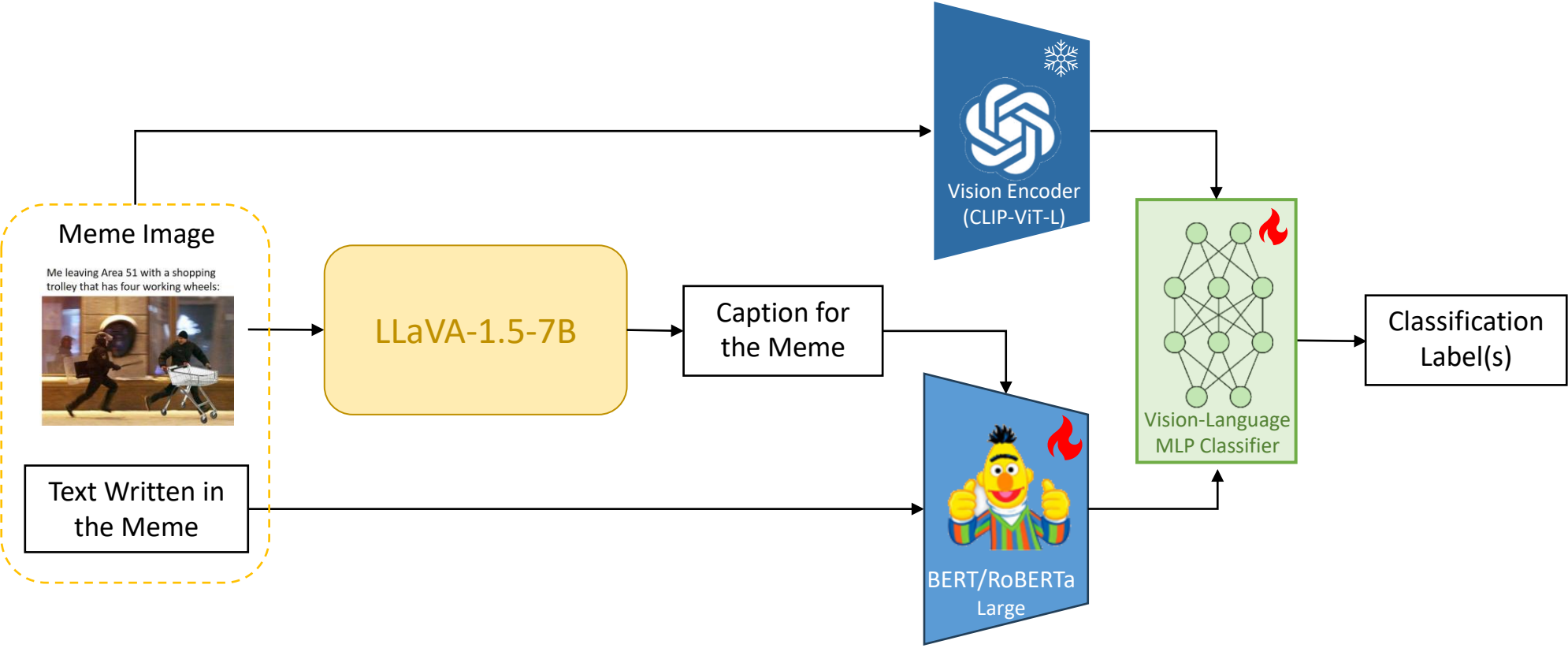
Vision Encoder  
(ResNet152)

BERT/roBERTa  
Large

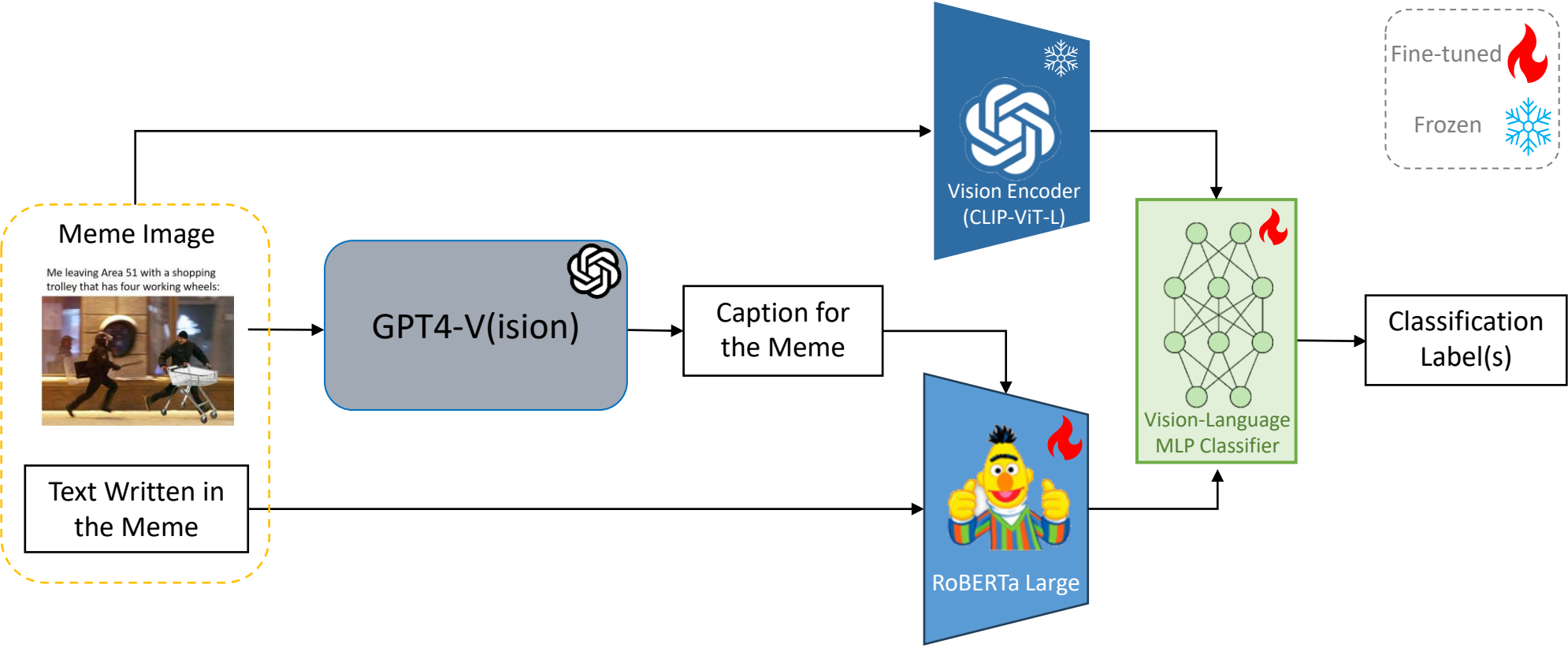
Vision-Language  
MLP Classifier

Classification  
Label(s)









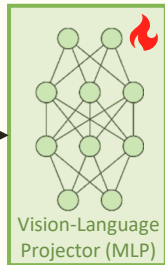
# MemeCap Dataset

Me leaving Area 51 with a shopping trolley that has four working wheels:





OCR

Meme  
Text



LLaVA-1.5-7B



Fine-tuned   
Frozen 

Caption for  
the Meme