

# Project Workflow

## Phase 1: Initialization

Load CLIP  
ViT-L/14

Load BLIP-2  
Captioner

Initialize  
LLM API

Create Class  
Prototypes

## Phase 2: Inference Pipeline

1. Input Image  
+ Hint

2. Domain  
Inference

3. Generate  
Prompts

4. CLIP  
Encoding

5. Compute  
Similarity

6. Caption  
Generation

7. LLM Reasoning

8. Final Prediction

## Phase 3: Adaptive Learning

Confidence  
> 0.15?

Yes

Update Prototype  
 $q \leftarrow (1-\alpha)q + \alpha v$

Skip Update

No

Return: Label + Confidence + Explanation