



# INSTAMED

CM GROUP 10

Kuber Budhija

Ankush Gupta

Pooja Nardeosingh

Nikita Bhatia

Priya Yadav

Medha Kashyap

Madhavi

Abuzar

# TABLE OF CONTENT

---

# InstaMed

- 01** Background
- 02** Data and Analysis
- 03** Hypothesis
- 04** Individual Contribution

# BACKGROUND

---

## **Healthcare Challenges:**

Accurate and timely diagnosis of organ-related issues is crucial for effective treatment.

Misdiagnosis and lack of proper tools often lead to delays in treatment.

## **Emerging AI Solutions:**

Advances in AI, particularly in natural language processing (NLP) and computer vision, offer innovative solutions for healthcare.

Multi-modal AI models like CLIP are bridging the gap between visual data and textual insights.

# BACKGROUND

---

## Why CLIP for Healthcare?

Multi-modal Capabilities: CLIP integrates visual and textual data, making it ideal for tasks requiring image classification and context understanding.

Versatility: By training on a diverse dataset, CLIP can generalize across various medical contexts, including organ identification and associated conditions.

## Motivation for the Project:

Simplify and enhance organ-specific diagnosis by leveraging AI for robust classification.

Provide users with an intuitive chatbot interface for further medical insights and recommendations, ensuring accessibility and ease of use.

# STORYBOARDING

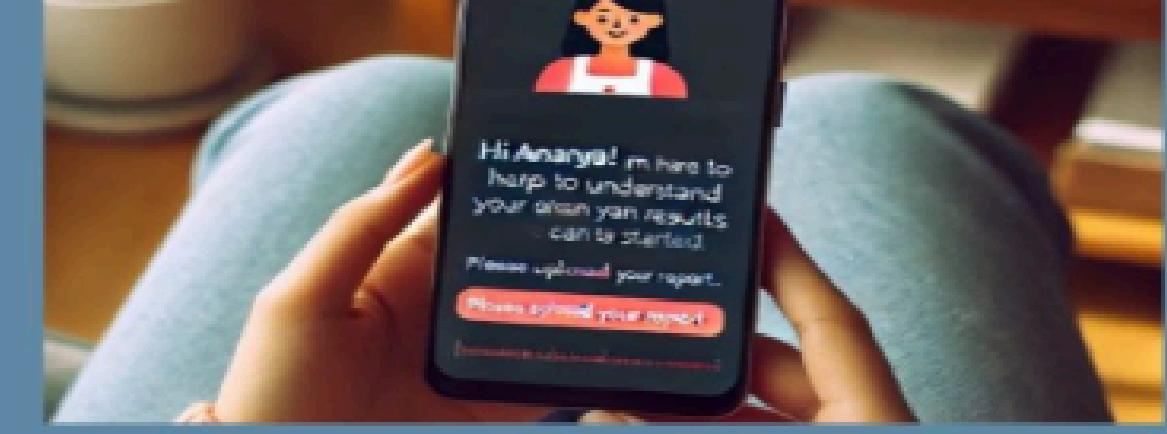
ANANYA FEELS UNWELL



SHE RECEIVES A SCAN ONLINE



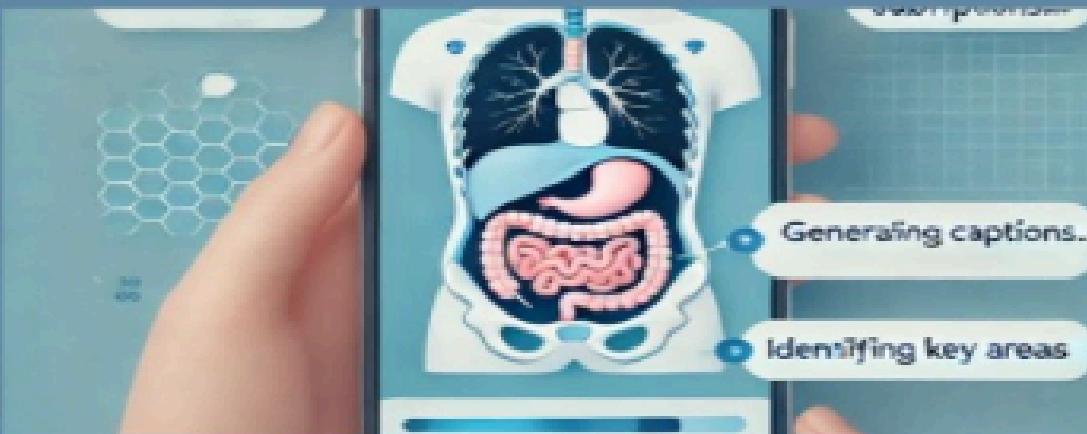
OPENS INSTAMED TO INTERPRET



UPLOADING THE SCAN



AI ANALYSIS



EXPLAINING THE SCAN



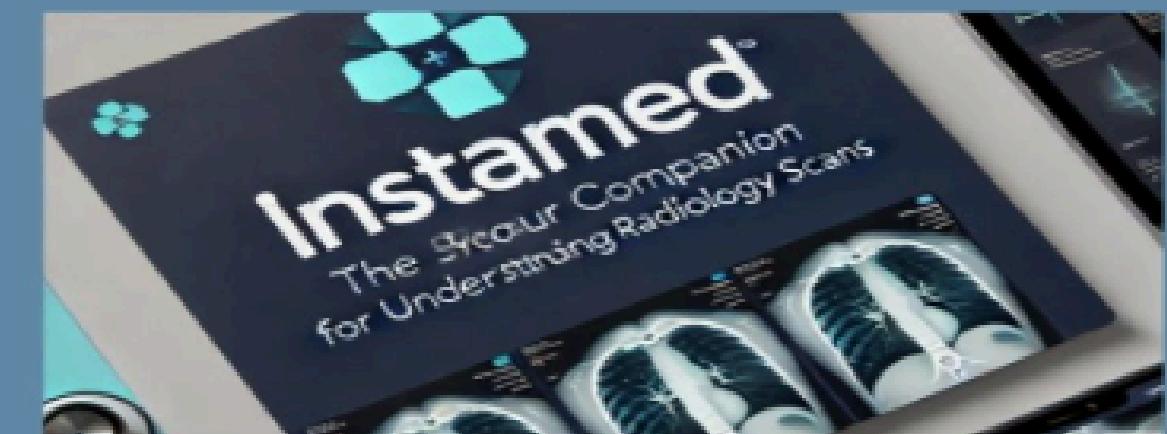
NEXT STEP AND RESOURCES



FEELING REASSURED



INSTAMED



# FHIR RESOURCES

FHIR (Fast Healthcare Interoperability Resources) is a standard for healthcare data exchange. For our model the following are the FHIR resources that align with storyboard and use case. The resources will include:

**Patient:** Representing the patient's demographic information.

**DiagnosticReport:** Representing the diagnosis scan report.

**Observation:** Containing details from the scan results (e.g., inflammation in the appendix).

**Communication:** For interactions between the patient and the InstaMed chatbot.

**CarePlan:** Suggesting next steps, such as saving results and following up with the doctor.



**Disclaimer: For the purpose this demo, we have updated the code to exclude using FHIR Resources, because HAPI server was unexpectedly down, here's a screenshot for the same**

# DATASET DESCRIPTION

**Dataset Archive:** Available on Open Science Framework

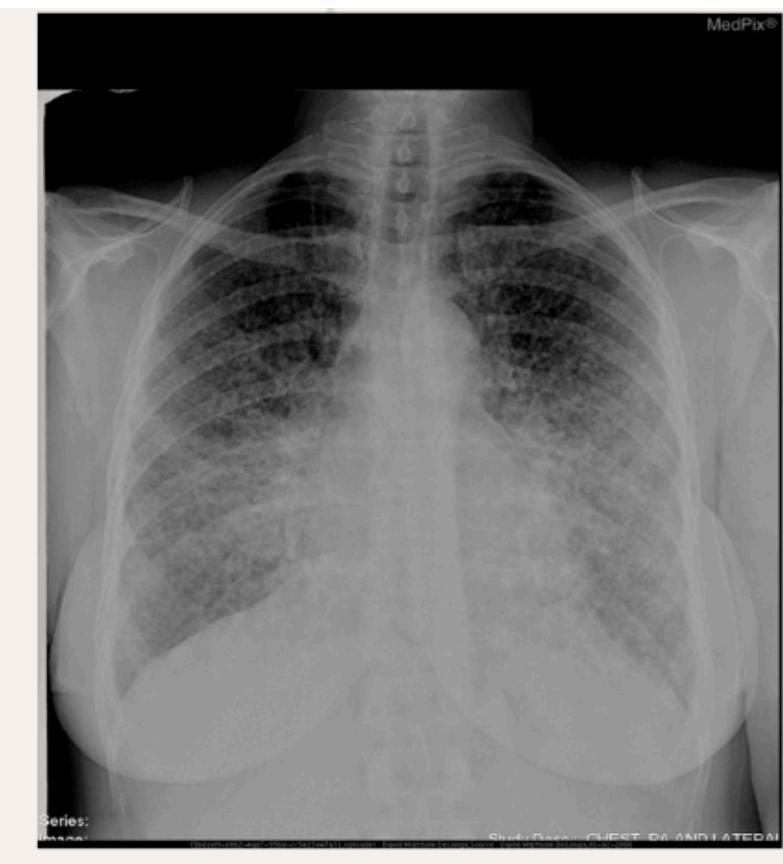
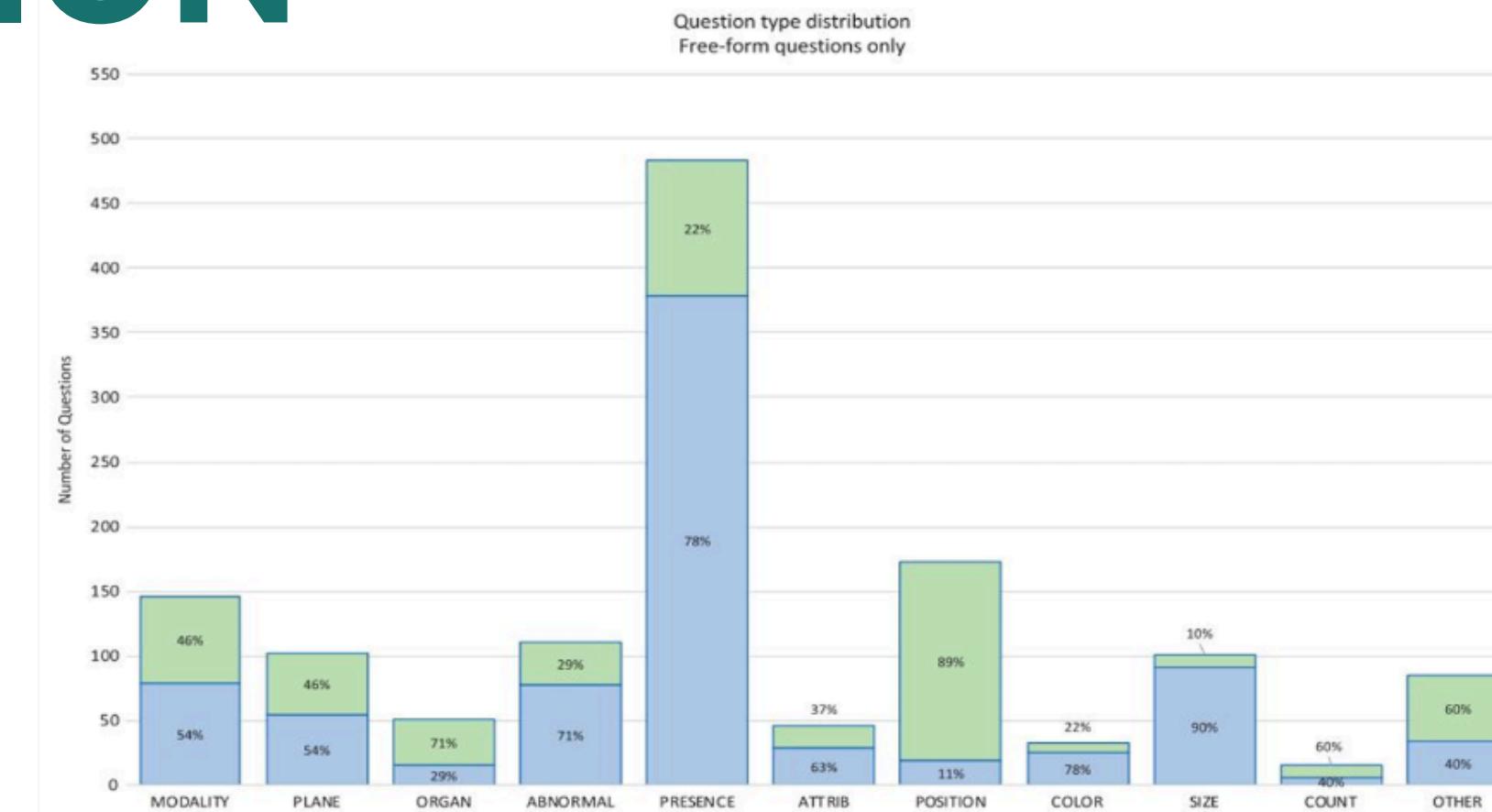
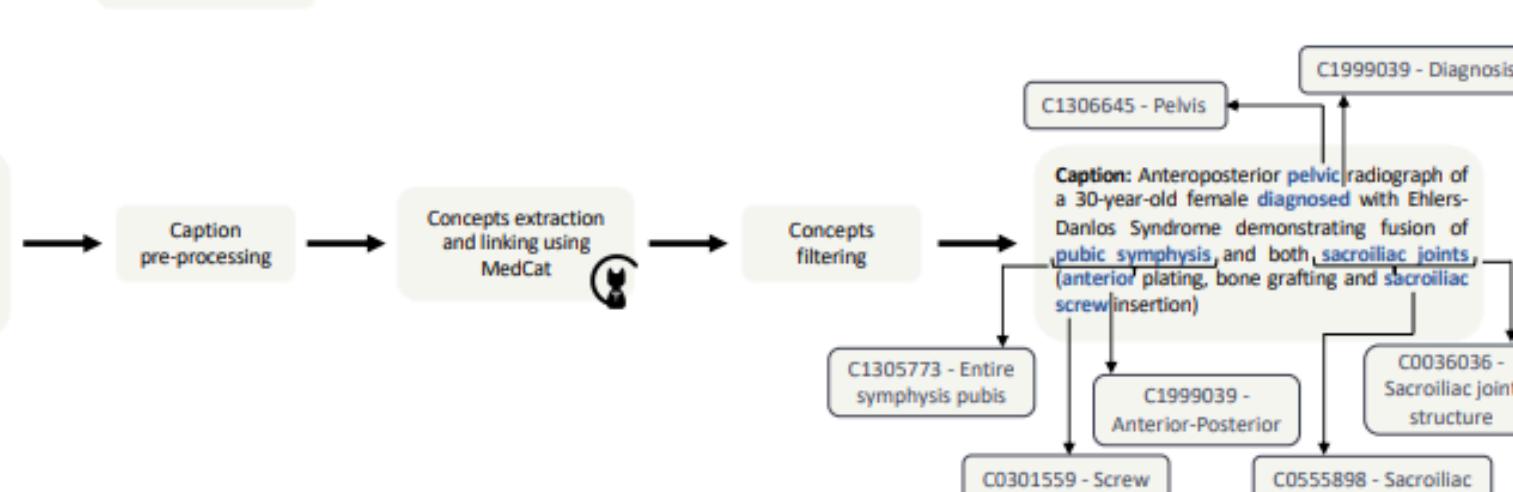
**Formats:** Question-answer pairs in JSON, XML, Excel.

**Dataset Overview:** 14 variables, 2,248 elements. Radiological

Images: 79,789 images (59,958 training) stored separately.



Caption: Anteroposterior pelvic radiograph of a 30-year-old female diagnosed with Ehlers-Danlos Syndrome demonstrating fusion of pubic symphysis and both sacroiliac joints (anterior plating, bone grafting and sacroiliac screw insertion)



Examples of images in the dataset

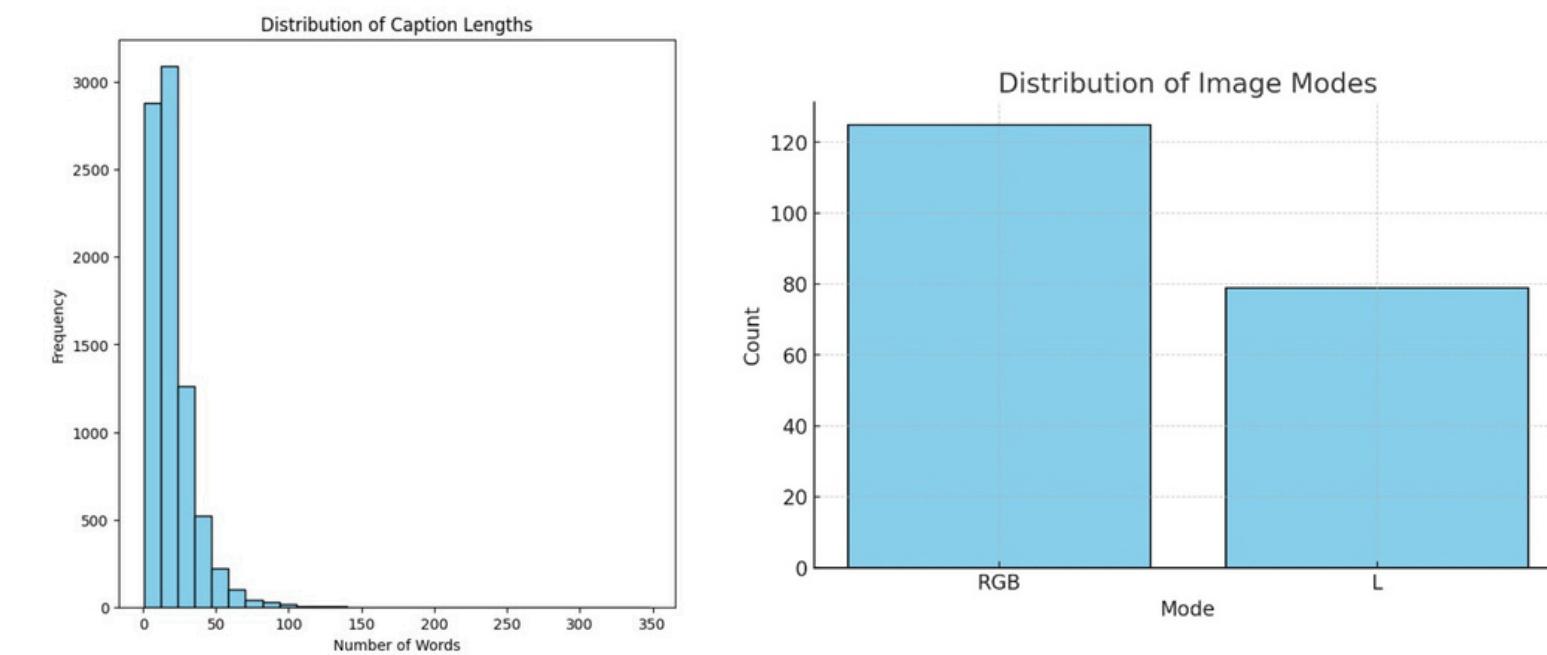
# EDA

EDA was conducted on the medical dataset containing images and corresponding captions. The analysis included examining basic caption properties like length, classifying captions into topics, and clustering images with their captions to identify patterns and relationships within the data.

Topic	Key Terms	Focus
1	left, right, artery, showing, view, pulmonary, vein, coronary, aortic, anterior	Anatomy of arteries and veins, focusing on left and right structures, pulmonary views.
2	radiograph, showing, left, right, bone, postoperative, lateral, panoramic, fracture, maxillary	Radiographic imaging of bones, including postoperative views, fractures, and maxillary structures.
3	CT, contrast, shows, showing, image, abdomen, abdominal, scan, small, axial	CT imaging of the abdomen, with contrast-enhanced scans and axial views.
4	image, MRI, magnetic, resonance, showing, imaging, axial, shows, left, sagittal	MRI imaging with axial and sagittal views, emphasizing magnetic resonance techniques.
5	tomography, showing, computed, chest, scan, left, right, CT, mass, multiple	Chest CT scans, computed tomography imaging, and detection of masses.

Table 1: Topics, Key Terms, and Focus Areas

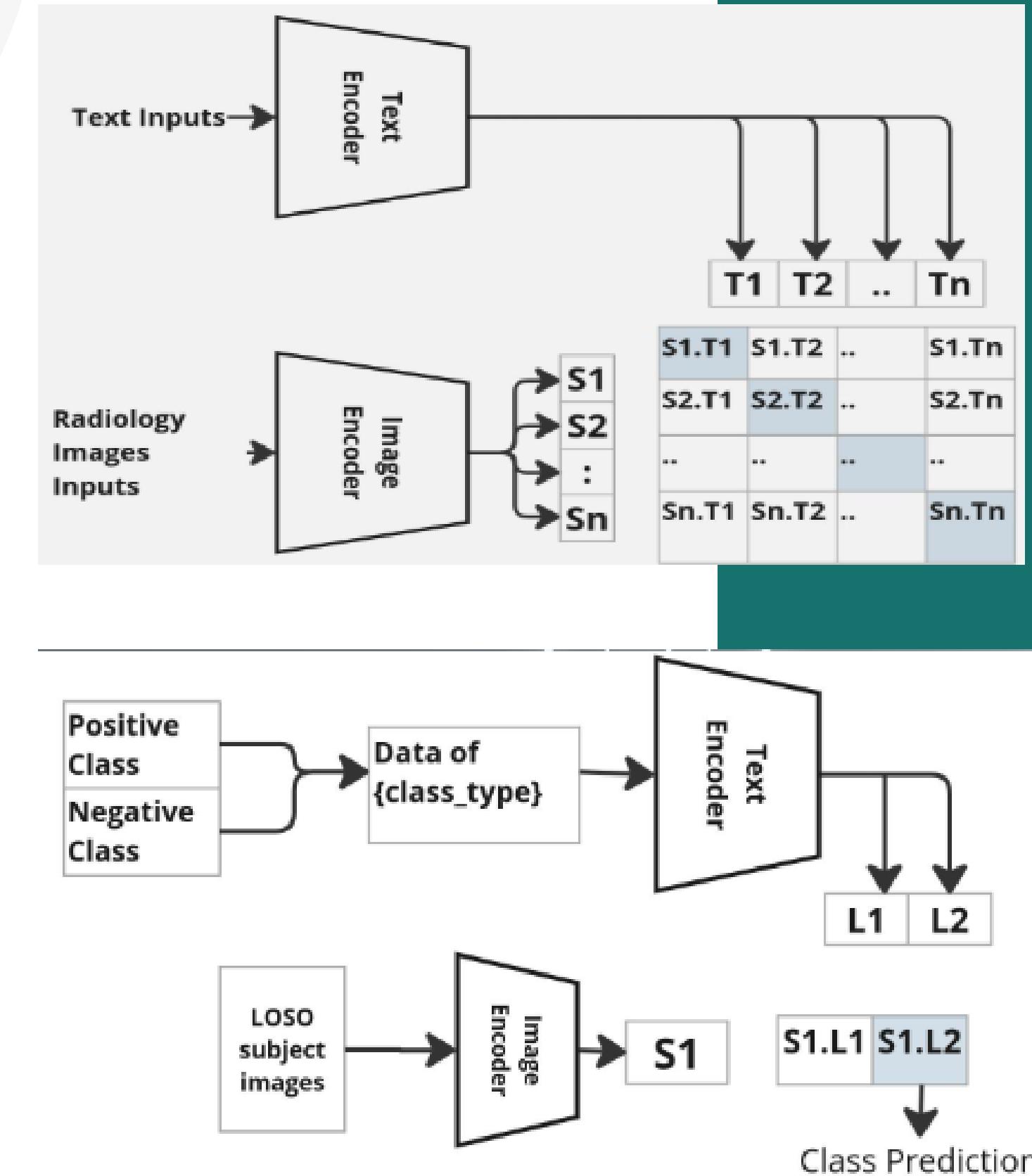
<b>mean</b>	<b>20.028606</b>
<b>std</b>	<b>15.462694</b>
<b>min</b>	<b>1.000000</b>
<b>25%</b>	<b>10.000000</b>
<b>50%</b>	<b>16.000000</b>
<b>75%</b>	<b>25.000000</b>
<b>max</b>	<b>348.000000</b>



Clustered Image and Caption Analysis		
Cluster	Details	Example Captions
0: MRI and Cyst-Related Terms	Captions describing MRI scans, cysts, and spinal hematomas.	<ul style="list-style-type: none"> <li>Magnetic resonance imaging pictures well-defined sellar-suprasellar heterogenous lesion and mildly hyperintense on T1-weighted (T1W) images and brightly hyperintense on T2-weighted (T2W) images suggestive of Rathke's cyst</li> <li>Transabdominal ultrasound in a 16-year-old adolescent. Notes: A hemorrhagic cyst consistent with a corpus luteum cyst diagnosed on transabdominal ultrasound in a 16-year-old adolescent who presented with acute abdominal pain. A hypoechoogenic cyst with an echogenic structure representing the blood clot is observed. In addition, the delicate "cobweb" is seen.</li> <li>Magnetic resonance imaging reveals a large acute spinal subarachnoid hematoma extending from C2 to T4 with spinal cord compression and numerous dilated anterior spinal arteries at the C6-T3 levels (white asterisk).</li> </ul>
		<p>Images</p>
1: Angiography, Stones, and Artery Occlusions	Captions related to angiography, stones, and artery occlusions.	<ul style="list-style-type: none"> <li>Angiography of the internal carotid artery, late arterial phase. (A) – venous drainage of the AVM, (B) – main arterial supplying vessel, arrow – AVM nidus.</li> <li>Preintervention MRCP revealing a nine-millimeter stone in the cystic duct (marked with arrow).</li> <li>MRI showing complete occlusion of proximal 4 cm of left subclavian artery with retrograde filling via the left vertebral.</li> </ul>

# HYPOTHESIS: CLIP

Fine-tuning the CLIP model on a radiological image dataset paired with corresponding expert-written captions will enhance its ability to understand domain-specific medical imagery and textual descriptions. This adaptation is expected to improve its performance in downstream radiological tasks such as disease classification, anomaly detection, and caption generation. By leveraging the pre-trained cross-modal alignment of CLIP and tailoring it to radiology, the fine-tuned model will demonstrate competitive or superior results compared to existing supervised radiological models, while retaining zero-shot capabilities for novel tasks in medical imaging.



Defining custom class for dataset

```
from torchvision import transforms
from torch.utils.data import Dataset

class CMDataSet(Dataset):
    def __init__(self, data):
        self.data = data
        self.transform = transforms.Compose([
            transforms.Resize(224, 224),
            transforms.ToTensor(),
            transforms.Normalize((0.48145466, 0.4578275, 0.40821073), (0.26862954, 0.26130258, 0.27577711))
        ])

    def __len__(self):
        return len(self.data)

    def __getitem__(self, idx):
        item = self.data[idx]
        image = item['image']
        text = item['caption']
        label = item['keywords'][0] + "_" + item['keywords'][1] + "_" + item['keywords'][2]
        return self.transform(image), label

from torch.utils.data import DataLoader
train_loader = DataLoader(CMDataSet(ds), batch_size=32, shuffle=True)
```

Modify the model to include a classifier for subcategories

```
class CLIPFineTuner(nn.Module):
    def __init__(self, model, num_classes):
        super(CLIPFineTuner, self).__init__()
        self.model = model
        self.classifier = nn.Linear(model.visual.output_dim, num_classes)

    def forward(self, x):
        with torch.no_grad():
            features = self.model.encode_image(x).float() # Convert to float32
        return self.classifier(features)

model_ft = CLIPFineTuner()
```

Define Loss function and optimizer and training the model

```
criterion = nn.CrossEntropyLoss()
optimizer = optim.Adam(model_ft.classifier.parameters(), lr=1e-4)

from tqdm import tqdm

num_epochs = 5
for epoch in range(num_epochs):
    model_ft.train()
    running_loss = 0.0
    pbar = tqdm(train_loader, desc=f'Epoch {epoch+1}/{num_epochs}, Loss: 0.0000')

    for images, labels in pbar:
        images, labels = images.to(device), labels.to(device)
        optimizer.zero_grad()
        outputs = model_ft(images)
        loss = criterion(outputs, labels)
        loss.backward()
        optimizer.step()

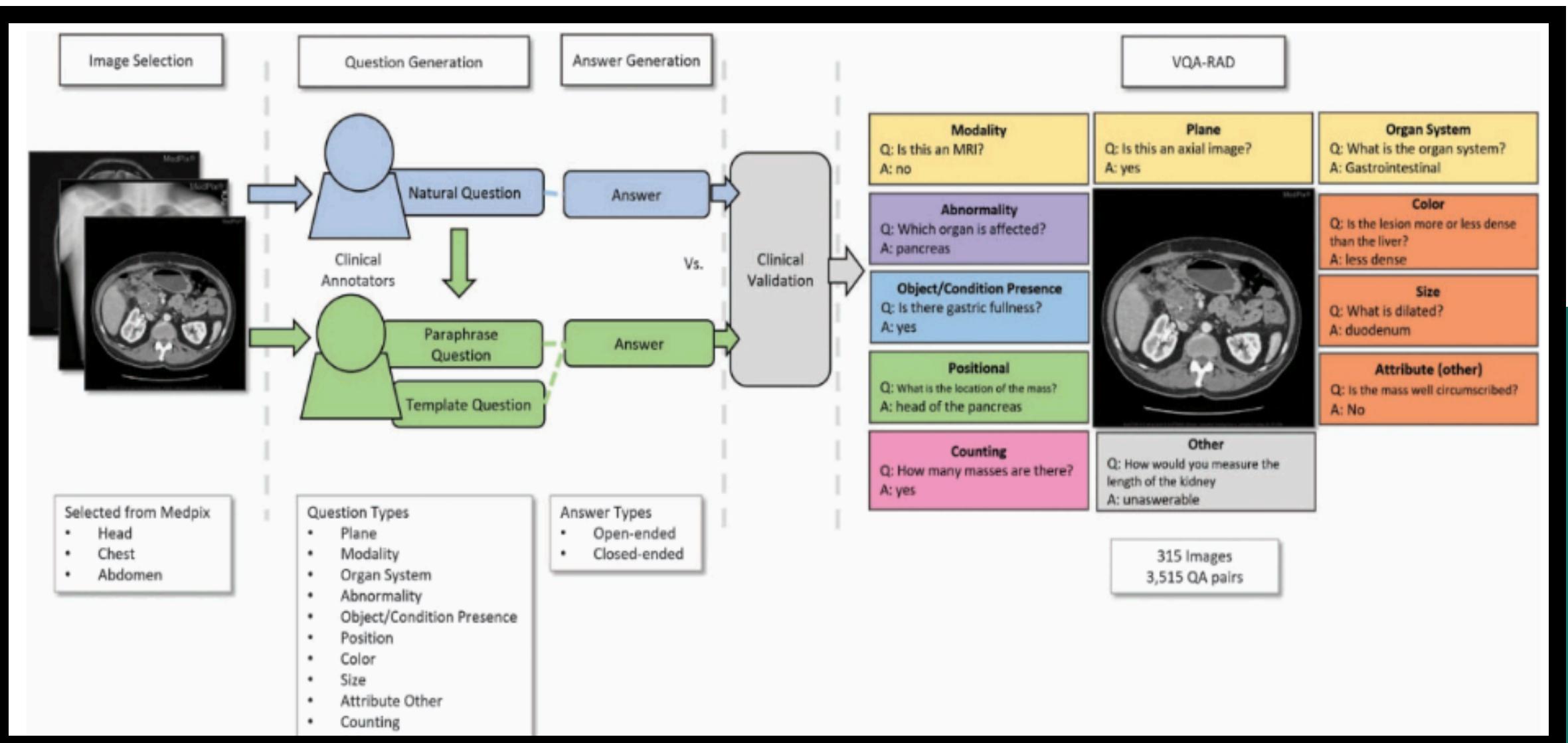
        running_loss += loss.item()
    pbar.set_description(f'Epoch [{epoch+1}/{num_epochs}], Loss: {running_loss/len(train_loader):.4f}')

print(f'Epoch [{epoch+1}/{num_epochs}], Loss: {running_loss/len(train_loader):.4f}')

torch.save(model_ft.state_dict(), 'clip-vit-base-CM')
```

# FINETUNING

The code snippets in the left side, shows how the finetuning has been done, over pre-existing CLIP, you can refer to the github link attached below for python file.



PYTHON NOTEBOOK FOR CODE OF TRAINING AND FINETUNING

# OUTPUTS OF FINE-TUNED CLIP ON DIFFERENT IMAGES



```
data = {
    'scan': {
        'type': 'Magnetic Resonance',
        'score': 0.9883
    },
    'region': {
        'type': 'Brain',
        'score': 0.9967
    },
    'abnormality': {
        'type': ['Metastatic Brain Tumors', 'Edema (Swelling)', 'Alzheimers Disease'],
        'score': [0.5531, 0.1417, 0.1188]
    }
}
```



```
data = {
    'scan': {
        'type': 'Computed Tomography Scan',
        'score': 0.9967
    },
    'region': {
        'type': 'Brain',
        'score': 0.9976
    },
    'abnormality': {
        'type': ['Parkinsons Disease', 'Edema (Swelling)', 'Skull Fractures'],
        'score': [0.4052, 0.2055, 0.1423]
    }
}
```

In the above two images and their corresponding outputs, we can see the results given by the CLIP in the data variable. It tells us the scan type, organ and abnormality with their probabilities.

# CHATBOT

WANNA TALK YOUR RESULTS?

```
class Gemini:
    def __init__(self):
        api_key = ""

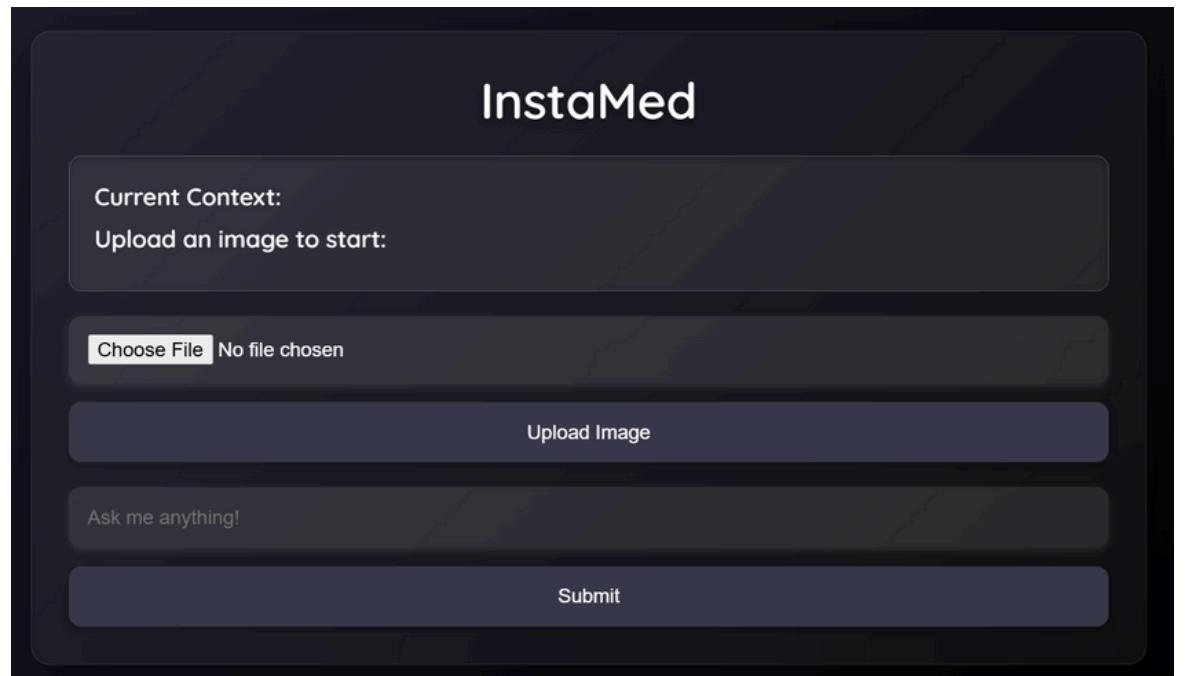
        genai.configure(api_key=api_key)
        generation_config = {
            "temperature": 0.9,
            "top_p": 1,
            "top_k": 1,
            "max_output_tokens": 2048,
        }

        safety_settings = [
            {
                "category": "HARM_CATEGORY_HARASSMENT",
                "threshold": "BLOCK_MEDIUM_AND_ABOVE"
            },
            {
                "category": "HARM_CATEGORY_HATE_SPEECH",
                "threshold": "BLOCK_MEDIUM_AND_ABOVE"
            },
            {
                "category": "HARM_CATEGORY_SEXUALLY_EXPLICIT",
                "threshold": "BLOCK_MEDIUM_AND_ABOVE"
            },
            {
                "category": "HARM_CATEGORY_DANGEROUS_CONTENT",
                "threshold": "BLOCK_MEDIUM_AND_ABOVE"
            },
        ]

        self.model = genai.GenerativeModel(model_name="gemini-pro", generation_config=generation_config, safety_settings=safety_settings)

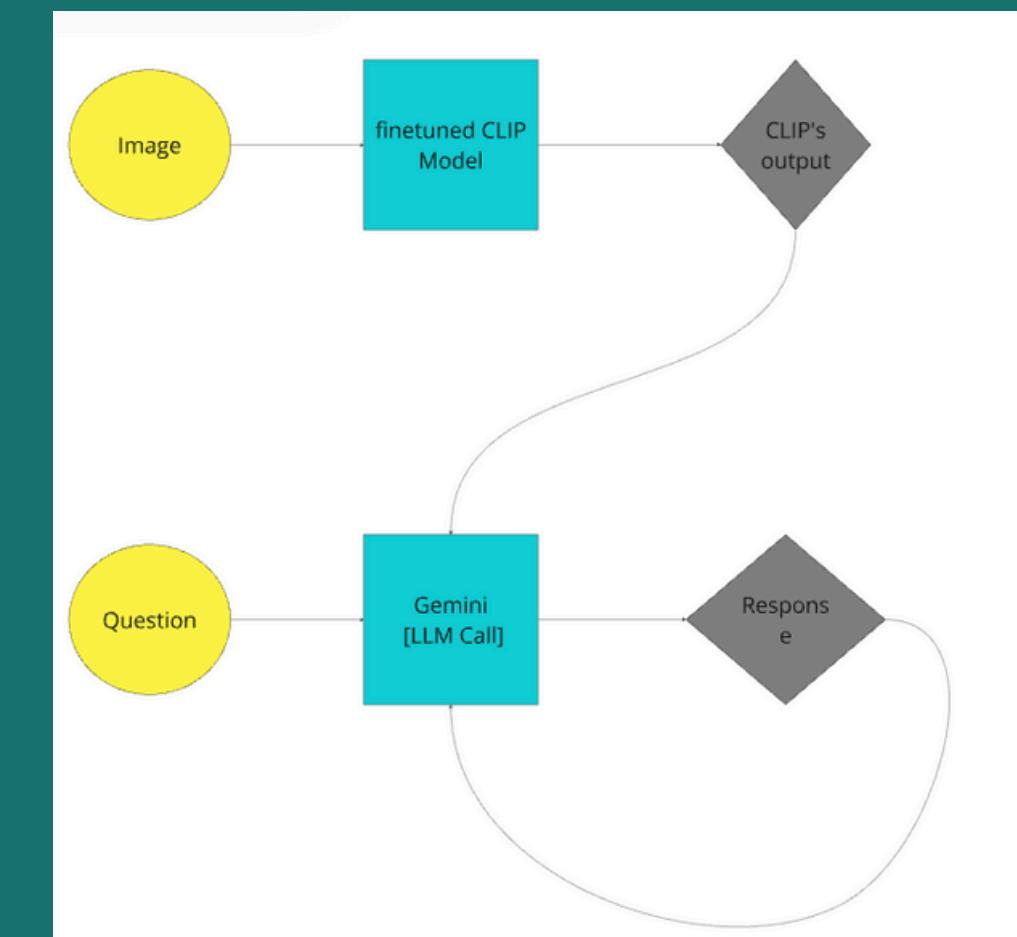
    def generate(self, prompt):
        prompt_parts = [prompt]

        response = self.model.generate_content(prompt_parts)
        return response.text
```



We also present you a chatbot, with which you can talk about anything you wanna know regarding your scan's analysis, your recommended food habits , precautions, next steps , we have gemini powered llm friend for you to talk it out

In the flow chart you can see, that CLIP's output, and previous questionnaire are being passed as context to LLM



The UI for the chatbot can be seen in the left hand side of slide, and we have attached code snippet explaining the class

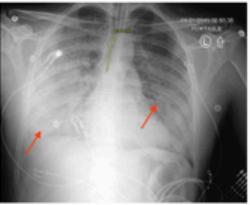
# RESULT COMPARISON

## Llava-Med

[PaperLink \(Microsoft @NeurIPS2023 \)](#)

Trained on [PMC-15M](#)

Visual input example, Biomedical image from [6]



Caption: Chest x-ray (CXR) on Day 2 of admission post-intubation (yellow line showing the level of the endotracheal tube). Rapidly worsening ground-glass opacities are seen throughout the lungs with relative subpleural sparing (red arrows)

Source: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7398722/>

What is shown in this image?

The image is a chest X-ray (CXR) that shows bilateral patchy infiltrates, which are areas of increased opacity in the lungs. These infiltrates can be indicative of various lung conditions, such as infections, inflammation, or other lung diseases.

Are there any devices implanted in this patient?

Yes, it appears that the patient has several devices implanted. These devices may include endotracheal tubes, central venous catheters, and/or pulmonary artery catheters. These devices are typically used for monitoring and managing various medical conditions, as well as for providing support during surgical procedures.

User  
LLaVA-Med

What is shown in this image?

The image shows a ribcage with wires coming out of it, indicating that it is likely a medical illustration or diagram.

User  
LLaVA-Med

Are there any devices implanted in this patient?

Yes, there are several devices protruding from the patient's chest, which are likely medical devices used for treatment or monitoring.

## CLIP

[PaperLink \(OpenAI \)](#)

Trained on [Coco Dataset](#)

What is shown in this image?

This is a chest X-ray image. It shows the thoracic cavity, including the lungs, ribs, heart, and other structures. The red arrows point to specific areas of interest, which could indicate abnormalities or findings that require further medical analysis.

Would you like me to help interpret this image in detail, keeping in mind I cannot diagnose or provide medical opinions, or do you want to share more context about this X-ray?

what should I do?

If this chest X-ray belongs to you or someone you're caring for, here's what you should do:

1. **Consult a Medical Professional:** Share the X-ray and any associated medical history or symptoms with a doctor or radiologist. They are qualified to interpret findings and provide accurate diagnoses.
2. **Prepare Relevant Information:**
  - Note any symptoms (e.g., chest pain, shortness of breath, cough, fever).
  - Share any past medical conditions or history of lung or heart issues.
3. **Follow-Up:** Based on the findings, the doctor may recommend:
  - Additional imaging or tests (e.g., CT scan, blood tests).
  - Specific treatments or medications if abnormalities are confirmed.
4. **Don't Panic:** X-rays can sometimes show artifacts or features that are not necessarily concerning. A professional review is essential.

If you'd like help formulating questions for the doctor or understanding medical terms, let me know!

## Finetuned CLIP

Fine-tuned on [Roco Dataset](#)

## InstaMed

Current Context:

The patient has an issue with Chest. The specific issue is Empyema.

what should i do, in such situation please suggest.

**\*\*Seek Medical Attention Immediately\*\*** Empyema is a serious medical condition that requires prompt medical attention. If you suspect you have empyema, seek medical attention immediately.  
**\*\*Treatment Options\*\*** Treatment typically involves:  
\* **Antibiotics:** To kill the underlying infection.  
\* **Thoracentesis:** Draining the pus and fluid from the pleural space.  
\* **Chest tube:** Insertion of a small tube into the pleural space to drain fluid and prevent re-accumulation.  
**Surgery:** In severe cases, surgery may be necessary to remove the infected tissue and drain the pus.  
**At Home Care** While receiving medical treatment, you can follow these steps to improve your comfort and recovery:  
\* **Rest and elevate:** Rest as much as possible and elevate your upper body to reduce pressure on your chest.  
\* **Apply heat:** Use a warm compress or heating pad on your chest to relieve pain and inflammation.  
\* **Take pain**

Choose File... New...

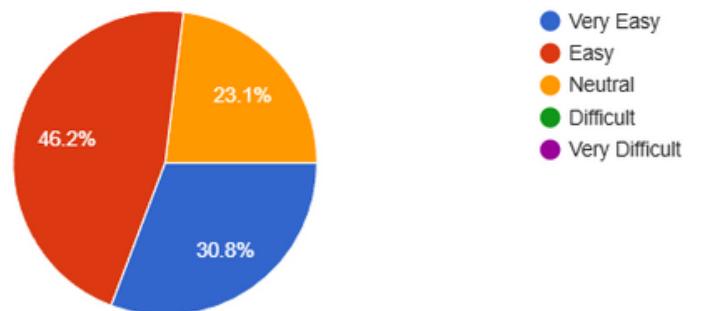
ta/Desktop/CM/project/code\_ankush.py  
Accuracy - Scan Type: 79.67  
Average Score 0.961

# USER STUDY

LET'S SEE HOW PEOPLE REACT TO IT

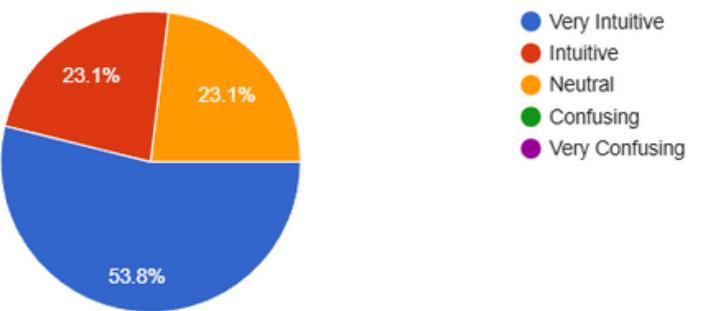
How easy was it to upload images and text to the model?

13 responses



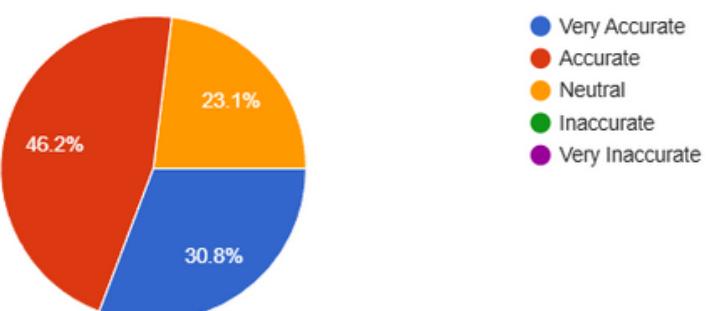
How intuitive was the user interface of the system?

13 responses



How accurately did the model identify diseases based on the inputs you provided?

13 responses



So we floated a google form among our friends and ask them for their opinion on the site, here are the results for the same, you can access google form [here](#)

## 1. Visibility of System Status

Feedback: The majority found the system's response time satisfactory and rated it "fast" or "very fast."

## 2. Match Between System and the Real World

Feedback: Many users described the interface as "very intuitive" or "intuitive." This indicates the tool aligns with users' expectations.

## 3. User Control and Freedom

Feedback: A user suggested adding a "chat saving option."

Overall, the feedback aligns well with the usability heuristics, showing strong areas of performance (e.g., ease of use, interface intuitiveness). Suggested enhancements include chat saving, user-friendly features, and robust error guidance to further refine the user experience.

# INDIVIDUAL CONTRIBUTION

- Background - Medha + Priya + Madhavi
- Textual Data cleaning - Medha + Priya
- EDA - Medha + Priya
- FHIR Resources - Kuber + Pooja
- Development - Kuber + Pooja
- Chatbot - Kuber + Pooja
- CLIP Fine Tuning - Ankush + Nikita
- Results - Ankush + Nikita
- User Study - Nikita + Abuzar
- Slides - All



# THANK YOU

● FOR YOUR NICE ATTENTION

