

# 2025 USA-NA-AIO Round 2, Problem 3, Part 1

USAAIO 

May 2025

## Problem 3 (100 points)

In this problem, you are asked to study **Contrastive Language-Image Pre-Training (CLIP)**, a powerful tool in multimodal AI.

 Topics

 My posts

⋮ More

▽ CATEGORIES

General

Site Feedback

 All categories

# Run code in this cell

....

DO NOT MAKE ANY CHANGE IN THIS CELL.

HINT: If something is not correctly installed, simply run this cell for few more  
....

!pip install datasets transformers



### WARNING !!!

- Beyond importing libraries/modules/classes/functions in the following cell, you are **NOT allowed to import anything else for the following purposes:**
  - **As a part of your final solution.** For instance, if a problem asks you to build a model without using sklearn but you use it, then you will not earn points.
  - **Temporarily import something to assist you to get a solution.** For instance, if a problem asks you to manually compute eigenvalues but you temporarily use np.linalg.eig to get an answer and then delete your code, then you violate the rule.

[Skip to main content](#)

**Rule of thumb:** Each part has its particular purpose to intentionally test you something. Do not attempt to find a shortcut to circumvent the rule.



# Run code in this cell

"""  
DO NOT MAKE ANY CHANGE IN THIS CELL.  
"""

```
from PIL import Image
import numpy as np
import matplotlib.pyplot as plt
from tqdm import tqdm

import torch
import torch.nn as nn
import torch.optim as optim
from torch.utils.data import DataLoader, Dataset

from transformers import BertTokenizer, BertModel, ViTModel
```

We will use flickr30k dataset to do image-language matching.

# Run code in this cell

"""  
DO NOT MAKE ANY CHANGE IN THIS CELL.  
"""

```
from datasets import load_dataset

dataset_train = load_dataset("USAAIO/2025-Round2-Problem3", split='train')
```

[Skip to main content](#)

## Part 1 (5 points, coding task)

**Do the following tasks to explore the properties of dataset\_train :**

1. dataset\_train is a list-like object. Print the number of elements in it.
2. Consider index idx = 2025 . Print the type of dataset\_train[idx] .
3. Print all keys in dataset\_train[idx] .
4. Name the value associated with the key image as image\_PIL . Print it.
5. Convert image\_PIL to a NumPy array object, called image\_np . Print image\_np and its shape.
6. Display this image by using plt.imshow .
7. Print the value associated with the key alt\_text . Print its type.

---

**USAAIO** **May 2025**

```
### WRITE YOUR SOLUTION HERE ###

print(len(dataset_train))

idx = 2025
print(type(dataset_train[idx]))
print(dataset_train[idx].keys())

image_PIL = dataset_train[idx]['image']
print(image_PIL)

image_np = np.array(image_PIL)
print(image_np)
print(image_np.shape)

plt.imshow(image_np)
plt.show()
```

**Skip to main content**

```
print(dataset_train[idx]['alt_text'])
print(type(dataset_train[idx]['alt_text']))

""" END OF THIS PART """
```

## ◆ Related topics

Topic	Replies	Activity
<a href="#">2025 USA-NA-AIO Round 2, Problem 3, Part 3</a>	1	May 2025
<a href="#">2025 USA-NA-AIO Round 2, Problem 3, Part 8</a>	1	May 2025
<a href="#">2025 USA-NA-AIO Round 2, Problem 3, Part 10</a>	1	May 2025
<a href="#">2025 USA-NA-AIO Round 2, Problem 3, Part 13</a>	1	May 2025
<a href="#">2025 USA-NA-AIO Round 2, Problem 3, Part 11</a>	1	May 2025