

2025 USA-NA-AIO Round 1, Problem 3, Part 8

Topics

More

CATEGORIES

General

Site Feedback

All categories

USAAIO 

Mar 2025

Part 8 (5 points, coding task)

Do the following tasks in this part.

1. Define a function called `my_train_test_split` that splits the whole dataset into the training component and the test/validation component.
 - The split is random
 - Inputs
 - `X` : A DataFrame object of features of all sample data.
 - `y` : A Series object of labels of all sample data.
 - `test_size` : It takes a value between 0 and 1 that denotes the fraction of samples used for testing. That is, the number of samples used for testing is `int(total number of samples * test_size)`.
 - Outputs
 - `X_train` : It keeps samples in `X` for training.
 - `X_test` : It keeps samples in `X` for testing.
 - `y_train` : It keeps samples in `y` for training.
 - `y_test` : It keeps samples in `y` for testing.
2. Call this function with inputs



[Skip to main content](#)

- o `X = X`
- o `y = y`
- o `test_size = 0.2`

3. Print object types and shapes of `X_train`, `X_test`, `y_train`, `y_test`.

USAAIO 

Mar 2025

```
### WRITE YOUR SOLUTION HERE ###
def my_train_test_split(X, y, test_size):
    num_samples = X.shape[0]
    num_test_samples = int(num_samples * test_size)
    indices = np.random.permutation(num_samples)
    test_indices = indices[:num_test_samples]
    train_indices = indices[num_test_samples:]

    X_train = X.iloc[train_indices]
    X_test = X.iloc[test_indices]
    y_train = y.iloc[train_indices]
    y_test = y.iloc[test_indices]

    return X_train, X_test, y_train, y_test

X_train, X_test, y_train, y_test = my_train_test_split(X, y, 0.2)

print(type(X_train))
print(type(X_test))
print(type(y_train))
print(type(y_test))

print(X_train.shape)
print(X_test.shape)
print(y_train.shape)
print(y_test.shape)
```

[Skip to main content](#)

"" END OF THIS PART """

◆ Related topics

| Topic | Replies | Activity |
|--|---------|--------------------------|
| 2025 USA-NA-AIO Round 2, Problem 3, Part 2 | 1 | May 2025 |
| 2025 USA-NA-AIO Round 1, Problem 3, Part 9 | 1 | Mar 2025 |
| 2025 USA-NA-AIO Round 1, Problem 2, Part 4 | 1 | Mar 2025 |
| 2025 USA-NA-AIO Round 2, Problem 1, Part 4 | 1 | May 2025 |
| 2025 USA-NA-AIO Round 1, Problem 3, Part 1 | 2 | Dec 2025 |

 Powered by Discourse