

How can I remove duplicate elements from a NumPy array?

Authored by
stats writer

June 26, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I remove duplicate elements from a NumPy array?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=154433>

Removing duplicate elements from a NumPy array refers to the process of eliminating any repeated values within the array. This can be achieved by using the `unique()` function, which returns a new array with only the unique elements present in the original array. This function is useful for data analysis and manipulation, as it helps in accurately representing data without any redundancy. In order to remove duplicate elements, the `unique()` function can be applied to the desired array, and the resulting array can be assigned to a new variable, effectively removing any duplicate values. This process ensures efficient and organized data handling, making it an essential step in data analysis using NumPy arrays.

Remove Duplicate Elements from NumPy Array

You can use the following methods to remove duplicate elements in NumPy:

Method 1: Remove Duplicate Elements from NumPy Array

```
new_data = np.unique(data)
```

Method 2: Remove Duplicate Rows from NumPy Matrix

```
new_data = np.unique(data, axis=0)
```

Method 3: Remove Duplicate Columns from NumPy Matrix

```
new_data = np.unique(data, axis=1)
```

The following examples show how to use each method in practice.

Example 1: Remove Duplicate Elements from NumPy Array

The following code shows how to remove duplicate elements from a NumPy array:

```
import numpy as np

#create NumPy array
data = np.array()

#create new array that removes duplicates
new_data = np.unique(data)

#view new array
print(new_data)
```

Notice that all duplicates have been removed from the NumPy array and only unique values remain.

Example 2: Remove Duplicate Rows from NumPy Matrix

The following code shows how to remove duplicate rows from a NumPy matrix:

```
import numpy as np

#create NumPy matrix
data = np.array(
,
,
])

#create new array that removes duplicate rows
new_data = np.unique(data, axis=0)
#view new matrix
print(new_data)

]
```

Notice that all duplicate rows have been removed from the NumPy matrix and only unique rows remain.

Example 3: Remove Duplicate Columns from NumPy Matrix

```
import numpy as np

#create NumPy matrix
data = np.array(
,
])
```

```
#create new matrix that removes duplicate columns  
new_data = np.unique(data, axis=1)
```

```
#view new matrix  
print(new_data)
```

```
]
```

Notice that all duplicate columns have been removed from the NumPy matrix and only unique columns remain.

The following tutorials explain how to perform other common tasks in NumPy: