How can I convert a Pandas DataFrame into a NumPy array? Can you provide some examples?

Authored by

stats writer

May 1, 2024

Pandas DataFrame and NumPy array are two commonly used data structures in data analysis and manipulation. While Pandas DataFrame provides a tabular structure with labeled rows and columns, NumPy array is a high-performance multidimensional array object. To convert a Pandas DataFrame into a NumPy array, the built-in function "to_numpy()" can be used. This function converts the DataFrame into a NumPy array while preserving the column names. Some examples of converting a Pandas DataFrame into a NumPy array are shown below:

Example 1:
df = pd.DataFrame({'A':, 'B':})
arr = df.to_numpy()
print(arr)

Output:

]

Example 2:
df = pd.DataFrame({'Name':, 'Age':})
arr = df.to_numpy()
print(arr)

Output:

]

Overall, converting a Pandas DataFrame into a NumPy array is a simple process using the "to_numpy()" function and can be useful for further data analysis and manipulation.

Convert Pandas DataFrame to NumPy Array (With Examples)

You can use the following syntax to convert a pandas DataFrame to a NumPy array:

df.to_numpy()

The following examples show how to use this syntax in practice.

Example 1: Convert DataFrame with Same Data Types

The following code shows how to convert a pandas DataFrame to a NumPy array when each of the columns in the DataFrame is the same data type:

```
import pandas as pd

#create data frame
df1 = pd.DataFrame({'rebounds': ,
'points': ,
'assists': })

#view data frame
print(df1)
```

```
rebounds points assists
0 7 5 11
1 7 7 8
2 8 7 10
3 13 9 6
4 7 12 6
5 4 9 5
```

```
#convert DataFrame to NumPy array
new = df1.to_numpy()

#view NumPy array
print(new)

]

#confirm that new is a NumPy array
print(type(new))

<class 'numpy.ndarray'>

#view data type
print(new.dtype)

int64
```

The Numpy array has a data type of int64 since each column in the original pandas DataFrame was an integer.

Example 2: Convert DataFrame with Mixed Data Types

The following code shows how to convert a pandas DataFrame to a NumPy array when the columns in the DataFrame are not all the same data type:

```
import pandas as pd

#create data frame
df2 = pd.DataFrame({'player': ,
'points': ,
'assists': })

#view data frame
print(df2)

player points assists
0 A 5 11
1 B 7 8
2 C 7 10
3 D 9 6
4 E 12 6
5 F 9 5

#convert DataFrame to NumPy array
new = df2.to_numpy()

#view NumPy array
print(new)

]
```

#confirm that *new* is a NumPy array
print(type(new))

<class 'numpy.ndarray'>

#view data type
print(new.dtype)

object

The Numpy array has a data type of object since not every column in the original pandas DataFrame was the same data type.

Example 3: Convert DataFrame & Set NA Values

The following code shows how to convert a pandas DataFrame to a NumPy array and specify the values to be set for any NA values in the original DataFrame:

import pandas as pd

#create data frame
df3 = pd.DataFrame({'player': ,
'points': ,
'assists': })

```
#view data frame
print(df3)

player points assists
0 A 5 11
1 B 7 8
2 <NA> <NA> 10
3 D 9 6
4 E <NA> 6
5 F 9 5

#convert DataFrame to NumPy array
new = df3.to_numpy(na_value='none')

#view NumPy array
print(new)

]

#confirm that new is a NumPy array
print(type(new))

<class 'numpy.ndarray'>

#view data type
print(new.dtype)
```

object