

/home/subhodeep/projects/mgproj/br/lightning/lightning/tests/test
dataset.py

```
1  import numpy as np
2  import scipy.sparse as sp
3
4  from sklearn.utils.testing import assert_array_equal
5  from sklearn.utils.testing import assert_array_almost_equal
6  from sklearn.utils.testing import assert_equal
7
8  from sklearn.datasets.samples_generator import make_classification
9  from sklearn.utils import check_random_state
10
11  from lightning.dataset_fast import ContiguousDataset
12  from lightning.dataset_fast import FortranDataset
13  from lightning.dataset_fast import CSRDataset
14  from lightning.dataset_fast import CSCDataset
15
16  # Create test datasets.
17  X, _ = make_classification(n_samples=20, n_features=100,
18                           n_informative=5, n_classes=2, random_state=0)
19  X2, _ = make_classification(n_samples=10, n_features=100,
20                             n_informative=5, n_classes=2, random_state=0)
21
22  # Sparsify datasets.
23  X[X < 0.3] = 0
24
25  X_csr = sp.csr_matrix(X)
26  X_csc = sp.csc_matrix(X)
27
28  rs = check_random_state(0)
29
30
31  def test_contiguous_get_row():
32      ind = np.arange(X.shape[1])
33      ds = ContiguousDataset(X)
34      for i in xrange(X.shape[0]):
35          indices, data, n_nz = ds.get_row(i)
36          assert_array_equal(indices, ind)
37          assert_array_equal(data, X[i])
38          assert_equal(n_nz, X.shape[1])
39
40
41  def test_csr_get_row():
42      ds = CSRDataset(X_csr)
```

```

43     for i in xrange(X.shape[0]):
44         indices, data, n_nz = ds.get_row(i)
45         for jj in xrange(n_nz):
46             j = indices[jj]
47             assert_equal(X[i, j], data[jj])
48
49
50 def test_fortran_get_column():
51     ind = np.arange(X.shape[0])
52     ds = FortranDataset(np.asfortranarray(X))
53     for j in xrange(X.shape[1]):
54         indices, data, n_nz = ds.get_column(j)
55         assert_array_equal(indices, ind)
56         assert_array_equal(data, X[:, j])
57         assert_equal(n_nz, X.shape[0])
58
59
60 def test_csc_get_column():
61     ds = CSCDataset(X_csc)
62     for j in xrange(X.shape[1]):
63         indices, data, n_nz = ds.get_column(j)
64         for ii in xrange(n_nz):
65             i = indices[ii]
66             assert_equal(X[i, j], data[ii])

```