

```
In [1]: from sklearn.datasets import fetch_rcv1
import numpy as np
dataset = fetch_rcv1(shuffle=True, random_state=1)
X = dataset.data[0:1000]
ym = dataset.target[0:1000]
topics = dataset.target_names
y = []

for i in range(1000):
    t = False
    for j in range(103):
        if (t==False) and (ym[i, j] == 1):
            y.append(topics[j])
            t=True

print(y[:5])
print(len(y))

['M14', 'GCAT', 'C18', 'C15', 'GCAT']
1000
```

```
In [ ]: y = np.array(y)
y = y.reshape(y.shape[0], 1)
```