

```
1 # Bootstraps 5000 samples for training set and rest to test set
2
3 import numpy as np
4
5 def bootstrap(X, y, n_train=5000):
6     """
7     Creates training and testing sets with 5000 training examples
8     sampled with replacement
9
10    parameters
11    -----
12    X: feature list
13    y: targets
14    n_train: number
15
16    returns:
17    -----
18    X_train, X_test, y_train, y_test numpy arrays
19    """
20
21    n_samples = X.shape[0]
22    train_inds = np.random.randint(n_samples, size=n_train)
23
24    X_train = np.take(X, train_inds, axis=0)
25    y_train = np.take(y, train_inds, axis=0)
26
27    X_test = np.take(X, [i for i in range(n_samples) if i not in train_inds],
28    axis=0)
29    y_test = np.take(y, [i for i in range(n_samples) if i not in train_inds],
30    axis=0)
31
32    return X_train, X_test, y_train, y_test
```