

```

1  #!/usr/bin/env python3
2  # -*- coding: utf-8 -*-
3  """
4  Example 3.1 Load the MNIST data set
5  Created for EMCH 504 at USC
6  @author: Austin Downey
7  """
8
9  import IPython as IP
10 IP.get_ipython().run_line_magic('reset', '-sf')
11
12 import numpy as np
13 import scipy as sp
14 import matplotlib as mpl
15 import matplotlib.pyplot as plt
16 import sklearn as sk
17 from sklearn import linear_model
18 from sklearn import datasets
19
20 plt.close('all')
21
22
23 %% Load your data
24
25 # this fetches "a" MNIST dataset from openml and loads it into your environment
26 # as a Bunch, a Dictionary-like object that exposes its keys as attributes.
27 mnist = sk.datasets.fetch_openml('mnist_784', as_frame=False, parser='auto')
28
29 # calling the DESCR key will return a description of the dataset
30 print(mnist['DESCR'])
31
32 # calling the data key will return an array with one row per instance and one
33 # column per feature where each features is a pixel, as defined in the key feature_names
34 X = mnist['data']
35
36
37 # calling the target key will return an array with the labels
38 Y = np.asarray(mnist['target'], dtype=int)
39
40 # Each image is 784 features or 28x28 pixels, however, the features must be reshaped
41 # into a 29x29 grid to make them into a digit, where the values represents one
42 # the intensity of one pixel, from 0 (white) to 255 (black).
43
44 digit_id = 35 # An OK 5
45 # digit_id = 0 # An odd 5
46 # digit_id = 100 # A bad 5
47
48
49 test_digit = X[digit_id,:]
50 digit_resaped = np.reshape(test_digit, (28,28))
51
52 # plot an image of the random pixel you picked above.
53 plt.figure()
54 plt.imshow(digit_resaped, cmap = mpl.cm.binary, interpolation="nearest")
55 plt.title('A "'+str(Y[digit_id])+'" digit from the MNIST dataset')
56 plt.xlabel('pixel column number')
57 plt.ylabel('pixel row number')
58
59

```