

DATA PROCESSING (SPRINT 1)

Team ID	PNT2022TMID26205
Project Name	A Novel Method for Handwritten Digit Recognition System

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from tensorflow.keras.datasets import mnist

(x_train, y_train), (x_test, y_test) = mnist.load_data()

Downloading data from https://storage.googleapis.com/tensorflow/tf-
keras-datasets/mnist.npz
11490434/11490434 [=====] - 2s 0us/step

x_train
array([[0, 0, 0, ..., 0, 0, 0],
       [0, 0, 0, ..., 0, 0, 0],
       [0, 0, 0, ..., 0, 0, 0],
       ...,
       [0, 0, 0, ..., 0, 0, 0],
       [0, 0, 0, ..., 0, 0, 0],
       [0, 0, 0, ..., 0, 0, 0]],

       [[0, 0, 0, ..., 0, 0, 0],
        [0, 0, 0, ..., 0, 0, 0],
        [0, 0, 0, ..., 0, 0, 0],
        ...,
        [0, 0, 0, ..., 0, 0, 0],
        [0, 0, 0, ..., 0, 0, 0],
        [0, 0, 0, ..., 0, 0, 0]],

       [[0, 0, 0, ..., 0, 0, 0],
        [0, 0, 0, ..., 0, 0, 0],
        [0, 0, 0, ..., 0, 0, 0],
        ...,
        [0, 0, 0, ..., 0, 0, 0],
        [0, 0, 0, ..., 0, 0, 0],
        [0, 0, 0, ..., 0, 0, 0]],

       ...,

       [[0, 0, 0, ..., 0, 0, 0],
        [0, 0, 0, ..., 0, 0, 0],
        [0, 0, 0, ..., 0, 0, 0],
        ...,
        [0, 0, 0, ..., 0, 0, 0],
        [0, 0, 0, ..., 0, 0, 0],
        [0, 0, 0, ..., 0, 0, 0]]]
```

```
[ [0, 0, 0, ..., 0, 0, 0],  
  [0, 0, 0, ..., 0, 0, 0],  
  [0, 0, 0, ..., 0, 0, 0],  
  ...,
```

array([[0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	
0,													
0,		0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	
		0,	0],										
0,	[0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	
0,		0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	
		0,	0],										
0,	[0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	
0,		0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	
		0,	0],										
0,	[0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	
0,		0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	
		0,	0],										
0,	[0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	
0,		0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	
		0,	0],										
3,	[0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	
		18,	18,	18,	126,	136,	175,	26,	166,	255,	247,	127,	0,

0,	0,	0],										
170,	[0,	0,	0,	0,	0,	0,	0,	0,	30,	36,	94, 154,
0,		253,	253,	253,	253,	253,	225,	172,	253,	242,	195,	64, 0,
	0,	0],										
253,	[0,	0,	0,	0,	0,	0,	0,	49,	238,	253,	253, 253,
0,		253,	253,	253,	253,	251,	93,	82,	82,	56,	39,	0, 0,
	0,	0],										
253,	[0,	0,	0,	0,	0,	0,	0,	18,	219,	253,	253, 253,
0,		253,	198,	182,	247,	241,	0,	0,	0,	0,	0,	0,
	0,	0],										
253,	[0,	0,	0,	0,	0,	0,	0,	0,	80,	156,	107, 253,
0,		205,	11,	0,	43,	154,	0,	0,	0,	0,	0,	0,
	0,	0],										
253,	[0,	0,	0,	0,	0,	0,	0,	0,	0,	14,	1, 154,
0,		90,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,
	0,	0],										
253,	[0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0, 139,
0,		190,	2,	0,	0,	0,	0,	0,	0,	0,	0,	0,
	0,	0],										
190,	[0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0, 11,
0,		253,	70,	0,	0,	0,	0,	0,	0,	0,	0,	0,
	0,	0],										
35,	[0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,
0,		241,	225,	160,	108,	1,	0,	0,	0,	0,	0,	0,
	0,	0],										
0,	[0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,
0,		81,	240,	253,	253,	119,	25,	0,	0,	0,	0,	0,
	0,	0],										
0,	[0,	0,	0,	0,	0,	0,	0,	0,	0,	0,	0,
		0,	45,	186,	253,	253,	150,	27,	0,	0,	0,	0,

[illegible]

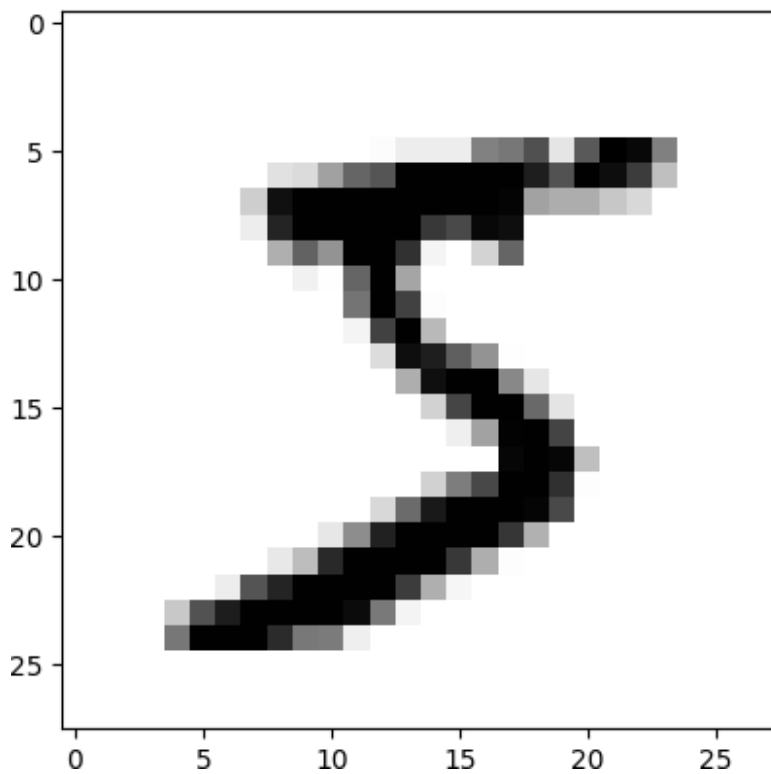
```

0,
    0,  0],
    [ 0,  0,  0,  0,  0,  0,  0,  0,  0,  0,  0,  0,  0,
    0,
        0,  0,  0,  0,  0,  0,  0,  0,  0,  0,  0,  0,
    0,
        0,  0],
    [ 0,  0,  0,  0,  0,  0,  0,  0,  0,  0,  0,  0,  0,
    0,
        0,  0,  0,  0,  0,  0,  0,  0,  0,  0,  0,  0,
    0,
        0,  0]], dtype=uint8)

```

```
plt.imshow(one_img, cmap='binary')
```

```
<matplotlib.image.AxesImage at 0x1c0f14903a0>
```



```
y_train
```

```
array([5, 0, 4, ..., 5, 6, 8], dtype=uint8)
```

```
from tensorflow.keras.utils import to_categorical
```

```
y_train.shape
```

```
(60000,)
```

```
y_example = to_categorical(y_train)
print(y_example, y_e)
```