

Snapshot of input and output

Dataset1

- input

```
train_x = read_file('/content/drive/MyDrive/ml_hw/Dataset1_train/train/X_train.xlsx')  
train_y = read_file('/content/drive/MyDrive/ml_hw/Dataset1_train/train/y_train.xlsx')  
test_x = read_file('/content/drive/MyDrive/ml_hw/Dataset1_test/X_test.xlsx')
```

- output

```
df = pd.DataFrame(np.array(preds))  
df.to_csv("Dataset1_pred.csv")
```

Dataset2

- input

```
# read_file(path): trianing_data features, trianing_data target, testing_data features  
train_x = read_file('/content/drive/MyDrive/ml_hw/Dataset2_train/Dataset2_train/X_train.xlsx')  
train_y = read_file('/content/drive/MyDrive/ml_hw/Dataset2_train/Dataset2_train/y_train.xlsx')  
test_x = read_file('/content/drive/MyDrive/ml_hw/Dataset2_test/X_test.xlsx')
```

- output

```
df = pd.DataFrame(np.array(preds))  
df.to_excel("Dataset2_pred.xlsx")
```