2)

a) OCR: - It is an optical character recognition system which uses a multilayer perceptron neural network classifier. It is a method that helps machines recognize texts. It has the advantage of being fast, easily trainable, and capable of creating arbitrary partitions of the input feature space.

b)

* Total number of lines in the file are 52152
* Fields which are specified in the file are id,letter, next\_id, word\_id, position, fold, and value of pixel in row i and column j.

This file contains handwritten words dataset.

c) Character visualizer is displaying the image of character or letter with a defined size and factor.

d)

* 50 data points
* 45
* 5
* 7

e) When the goal is reached or when maximum number of train epochs is reached.

f)

* Trial #1:
  + Time Elapsed: - 1 minutes and 10 seconds
  + Epochs ran: - 10000
  + Final Error: -0.04741649312598416
  + Predictions: -

Original: o

Predicted: o

Original: m

Predicted: n

Original: m

Predicted: n

Original: a

Predicted: n

Original: n

Predicted: n

* Trial #2:
  + Time Elapsed: - 1 minutes and 14 seconds
  + Epochs ran: - 10000
  + Final Error: - 0.2738007280206442
  + Predictions: -

Original: o

Predicted: o

Original: m

Predicted: n

Original: m

Predicted: m

Original: a

Predicted: n

Original: n

Predicted: n

g)

* Trial #1:
  + Time Elapsed: - 2 minutes and 25 seconds
  + Epochs ran: - 10000
  + Final Error: - 2.7441624872287145
  + Predictions: -

Original: o

Predicted: o

Original: m

Predicted: m

Original: m

Predicted: m

Original: a

Predicted: o

Original: n

Predicted: a

Original: d

Predicted: d

Original: i

Predicted: i

Original: n

Predicted: n

Original: g

Predicted: g

Original: o

Predicted: o

* Trial #2:
  + Time Elapsed: - 2 minutes and 26 seconds
  + Epochs ran: - 10000
  + Final Error: - 3.6338331206648826
  + Predictions: -

Original: o

Predicted: o

Original: m

Predicted: m

Original: m

Predicted: n

Original: a

Predicted: o

Original: n

Predicted: n

Original: d

Predicted: d

Original: i

Predicted: i

Original: n

Predicted: m

Original: g

Predicted: n

Original: o

Predicted: o

h)

When number of data points were increased time taken by the program increased and Error also increased. It seems with less data points, it has less error. More test by changing Epoch number can be done which might have better results.