

# Back Propagation on Feed Forward Neural Networks

## Report and Observations

<b>Group 12:</b>	Sahil Jindal	110020043
	Rohan Gyani	110040001
	Abhishek Gupta	110040067
	Mridul Ravi Jain	110040083

# Assignment Specifications

- Implement back propagation (BP) on feed forward neural n/w (FFNN). Give FFNNs for:
  - *2-input XOR*
  - *2-input NAND*
  - *5-input palindrome*
  - *5-input majority*
  - *5-input parity*
  - *Digit recogniser*
- Choose the learning rate judiciously.
- Study convergence time, local minima, saturation, effect of initialization, effect of learning rate and momentum factor.
- 1 and 0 decisions are based on the output being above the high water mark or being below the low water mark.

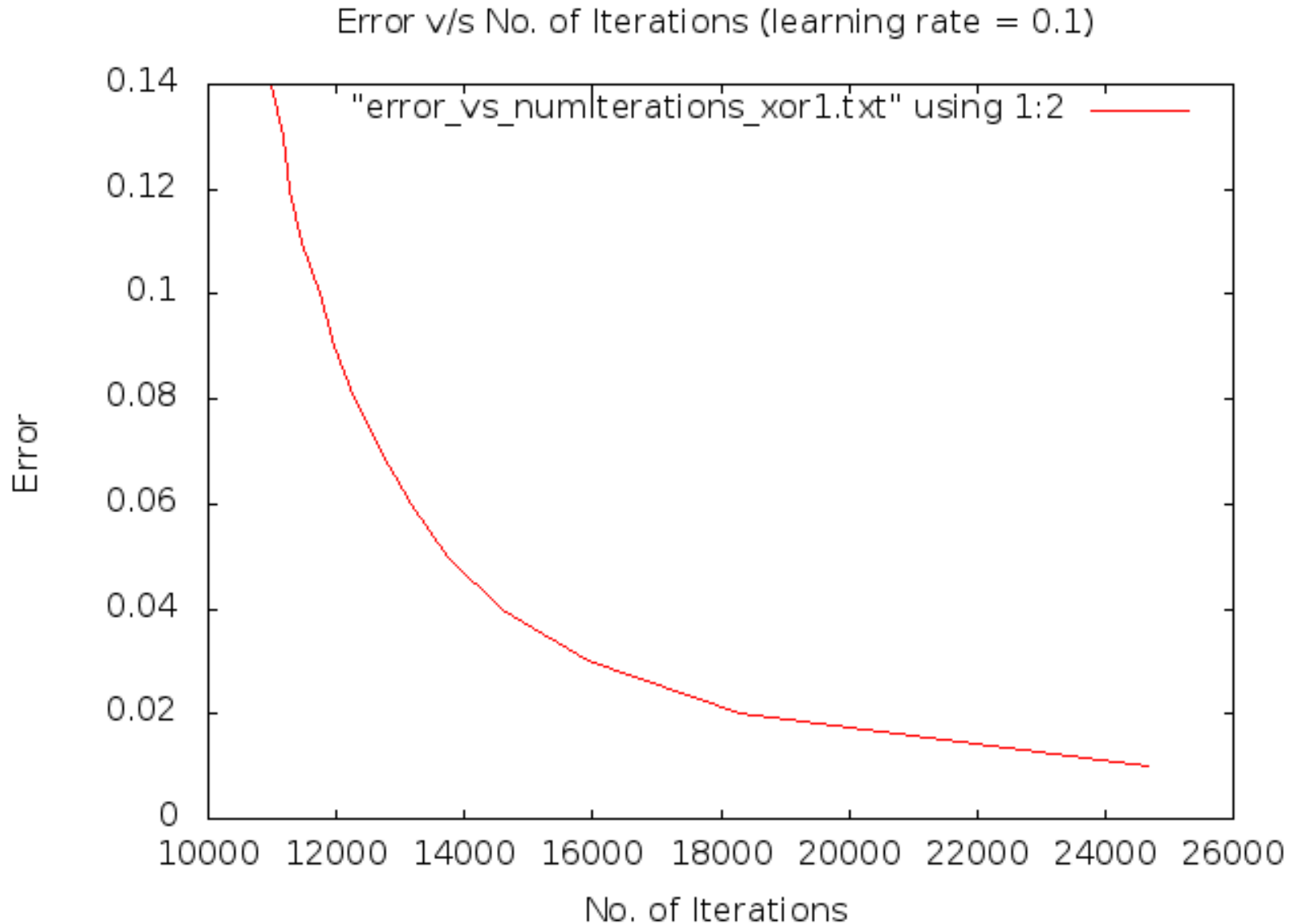
# FFNN Pseudo Code

- Apply the inputs to the network and work out the output.
- Calculate errors of output neurons and then calculate Total Sum of Squares Error
- while(Total Sum of Squares Error > Threshold)
  - Change output layer weights
  - Calculate errors for hidden layer neurons
  - Change hidden layer weights(backpropagation)
  - Apply the inputs to the network and work out the output.
  - Calculate errors of output neurons and then calculate Total Sum of Squares Error

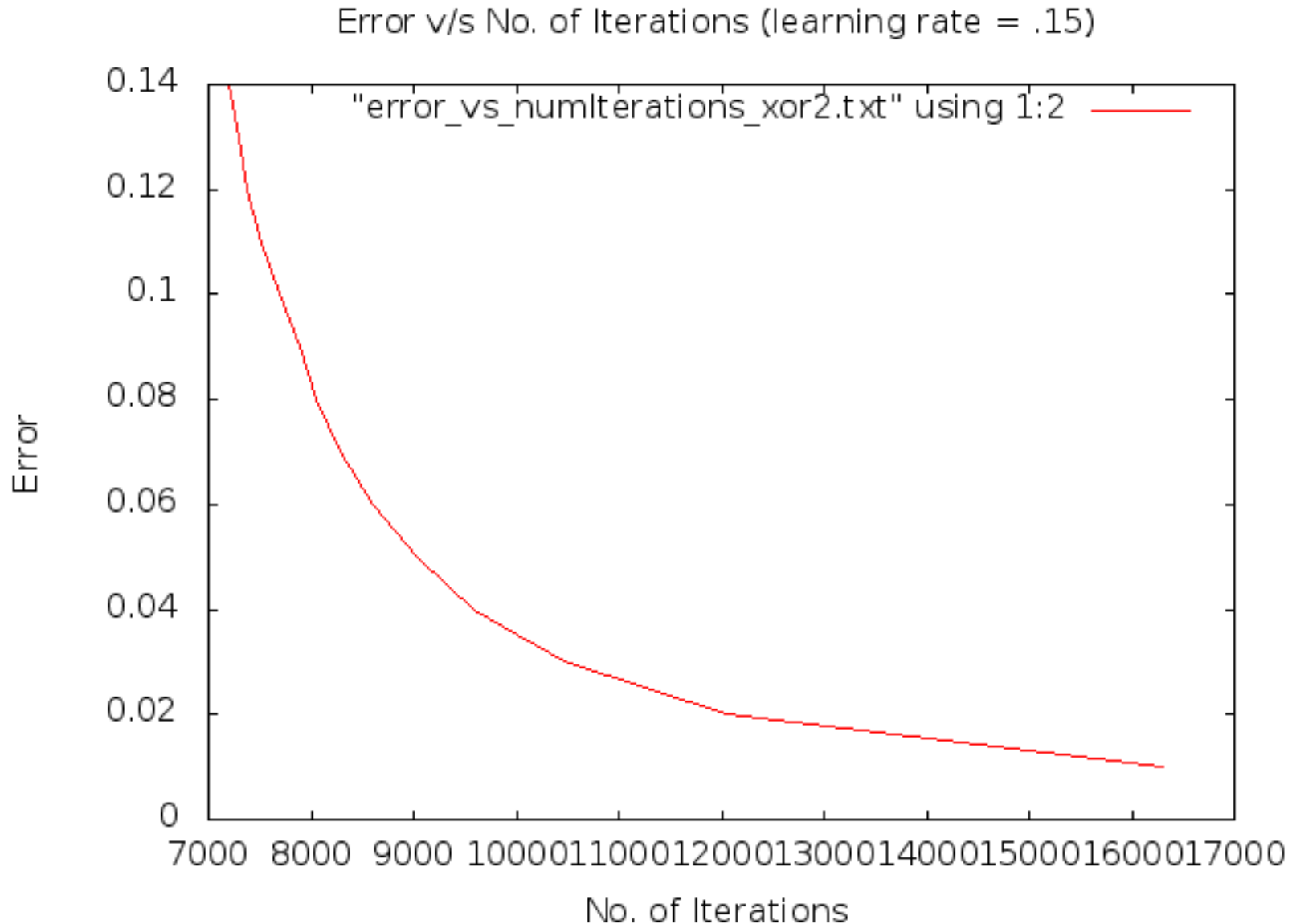
# Contents

- Error v/s No of Iterations
  - Effect of Learning Rate
  - Effect of Momentum factor
- Graphs for
  - *2-input XOR*
  - *2-input NAND*
  - *5-input palindrome*
  - *5-input majority*
  - *5-input parity*
  - *Digit recogniser*
- *Functionality of Hidden Layer Neurons*
  - *2-input XOR*
  - *3-input Palindrome*
  - *5-input Palindrome*

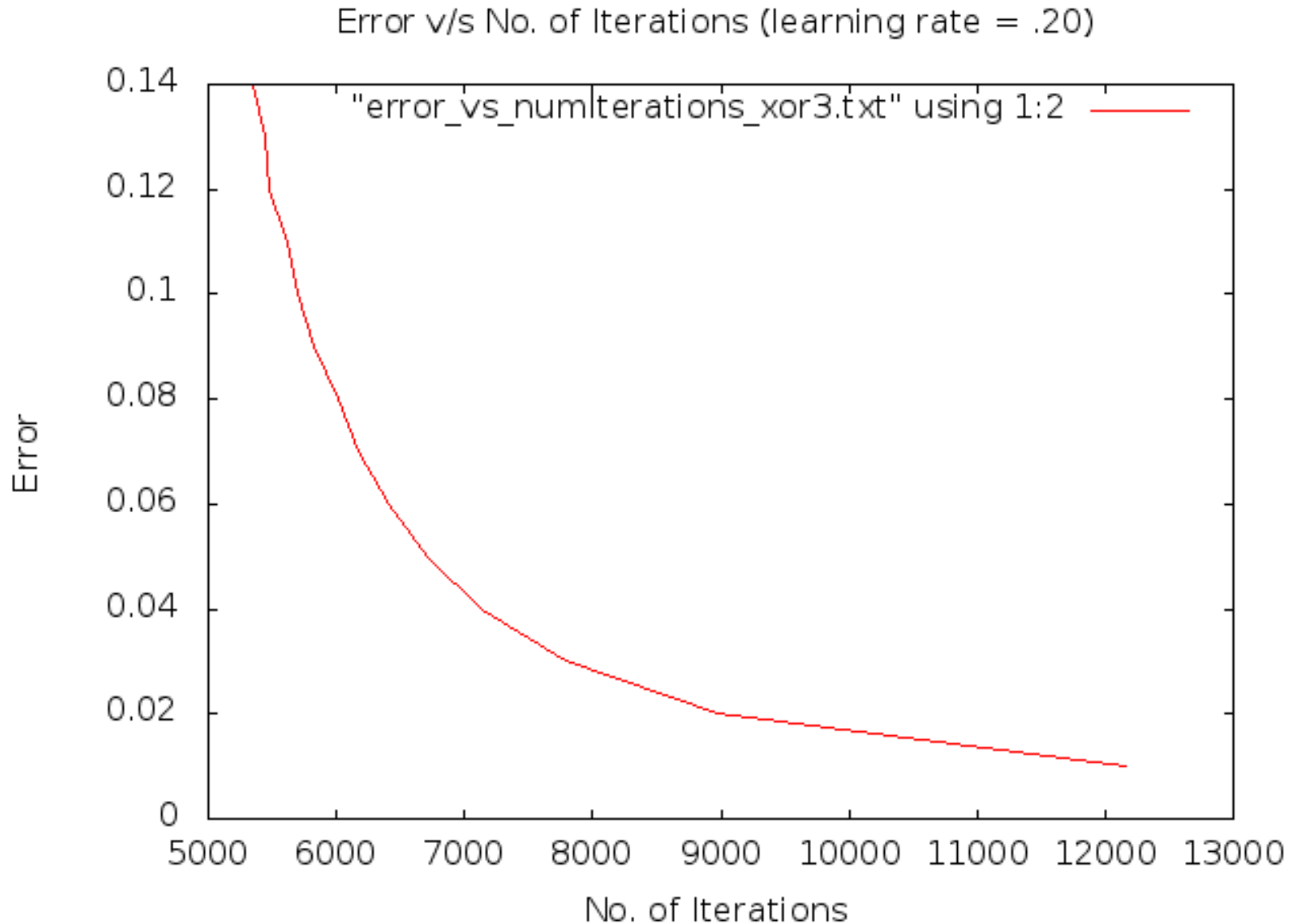
# Effect of Learning Rate (XOR)



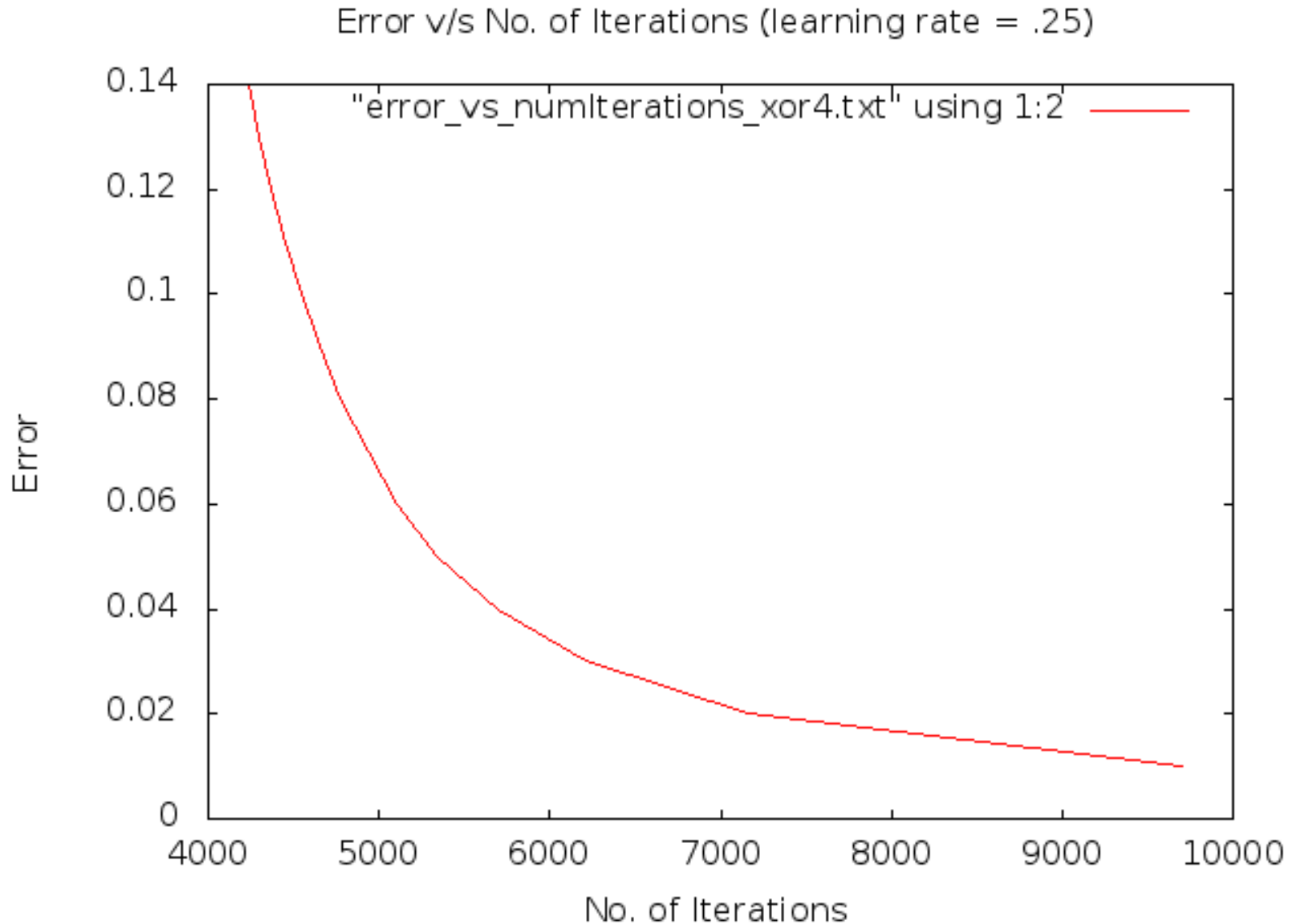
# Effect of Learning Rate (XOR)



# Effect of Learning Rate (XOR)

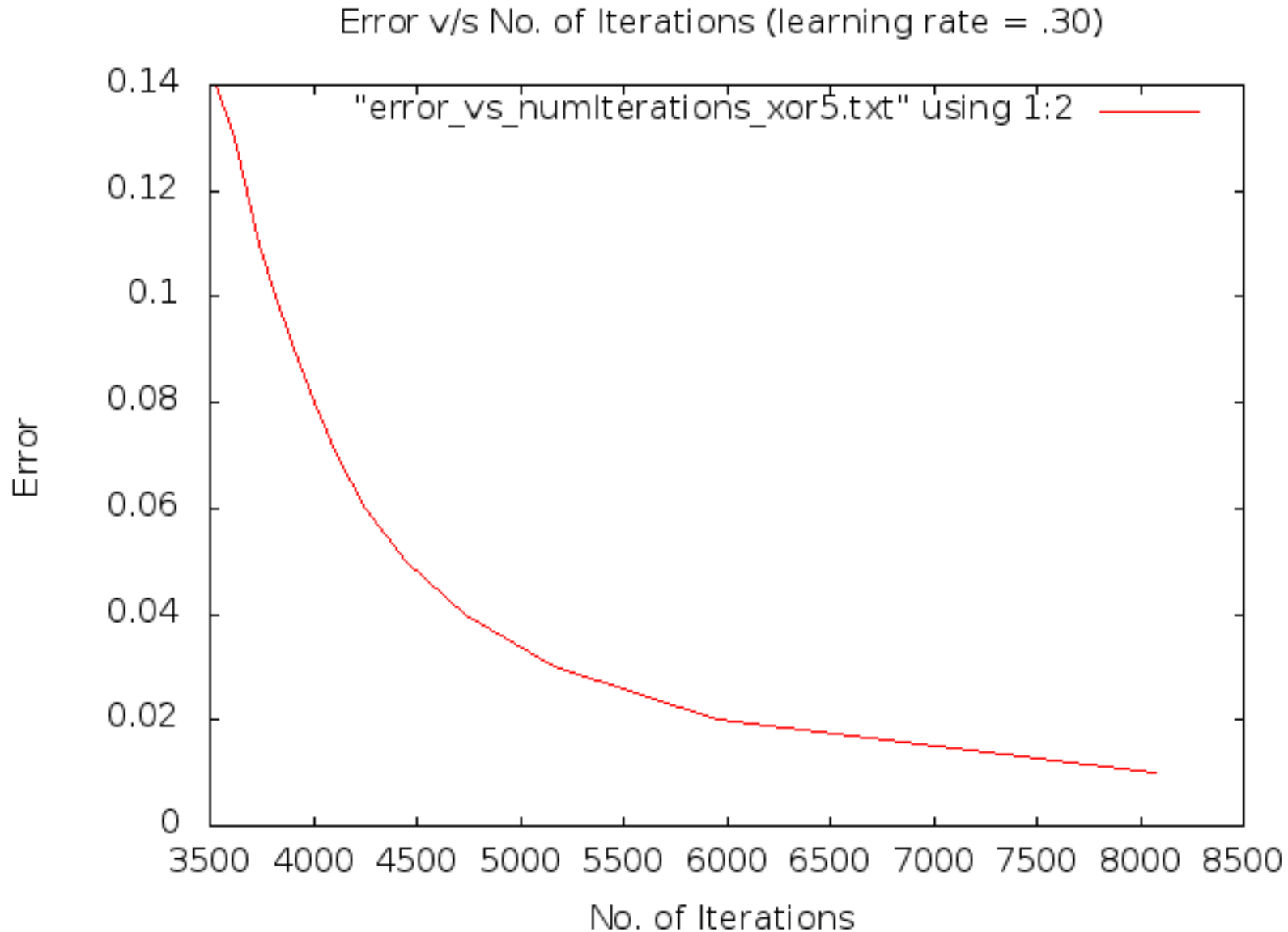


# Effect of Learning Rate (XOR)

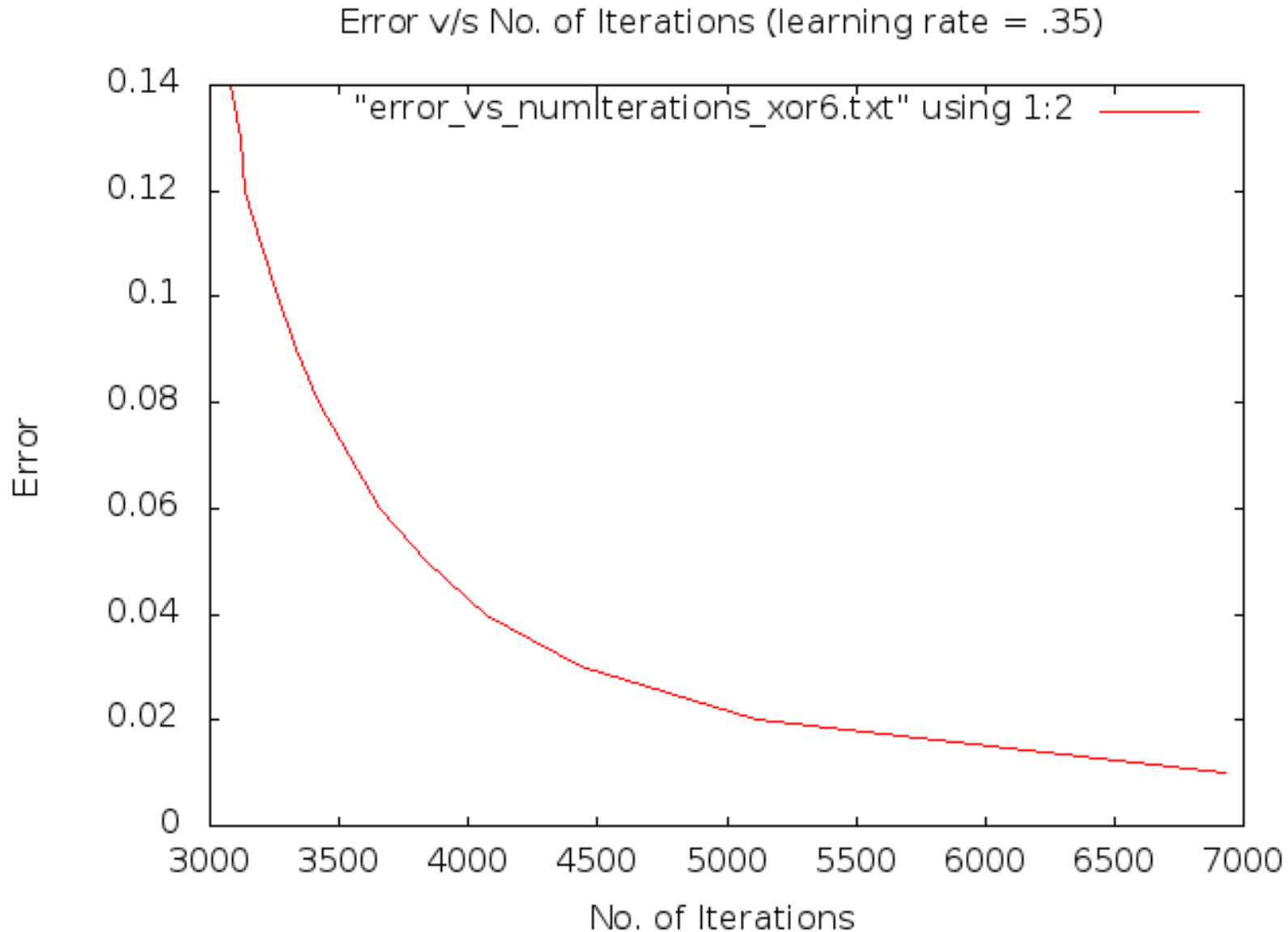




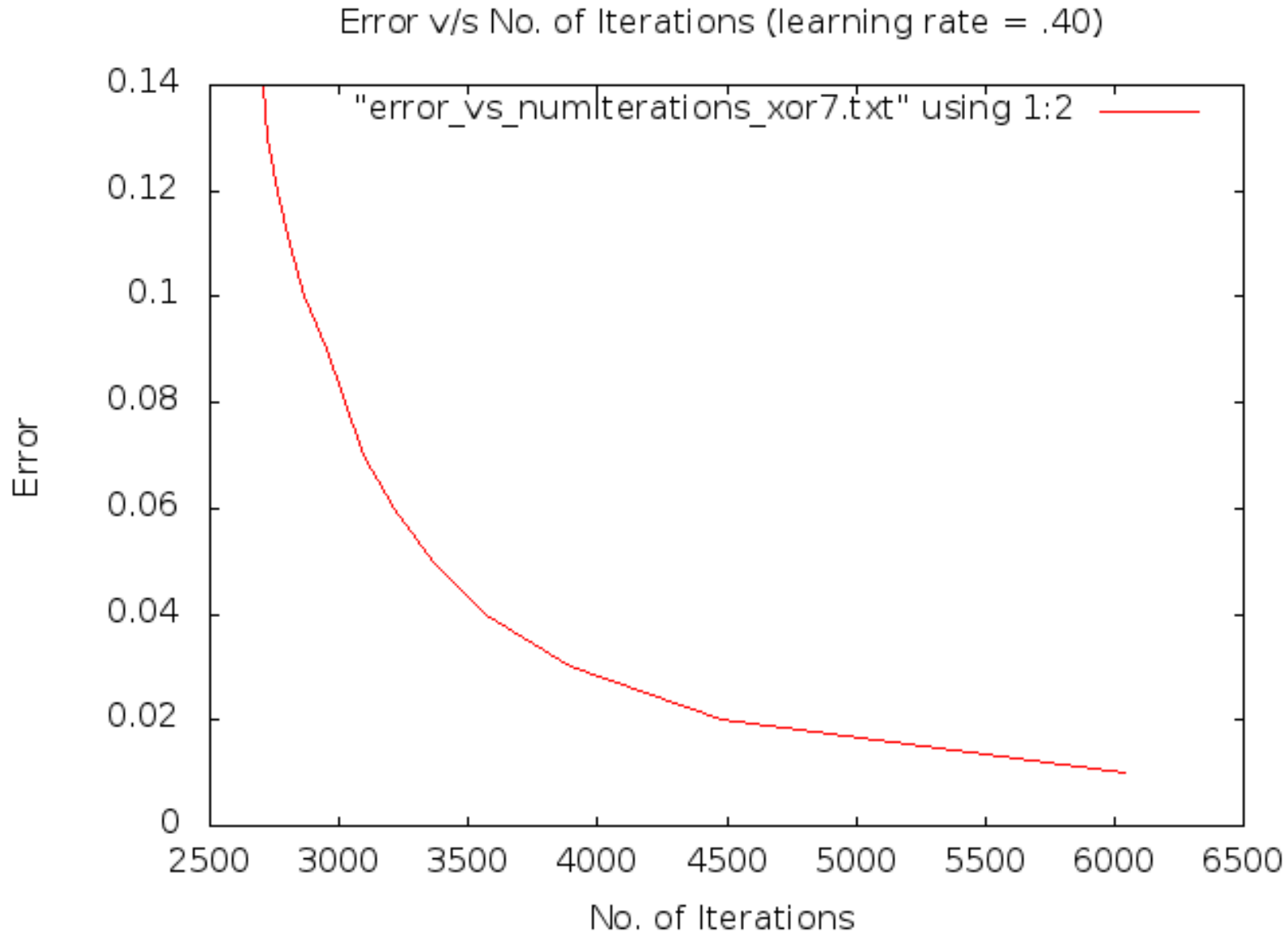
# Effect of Learning Rate (XOR)



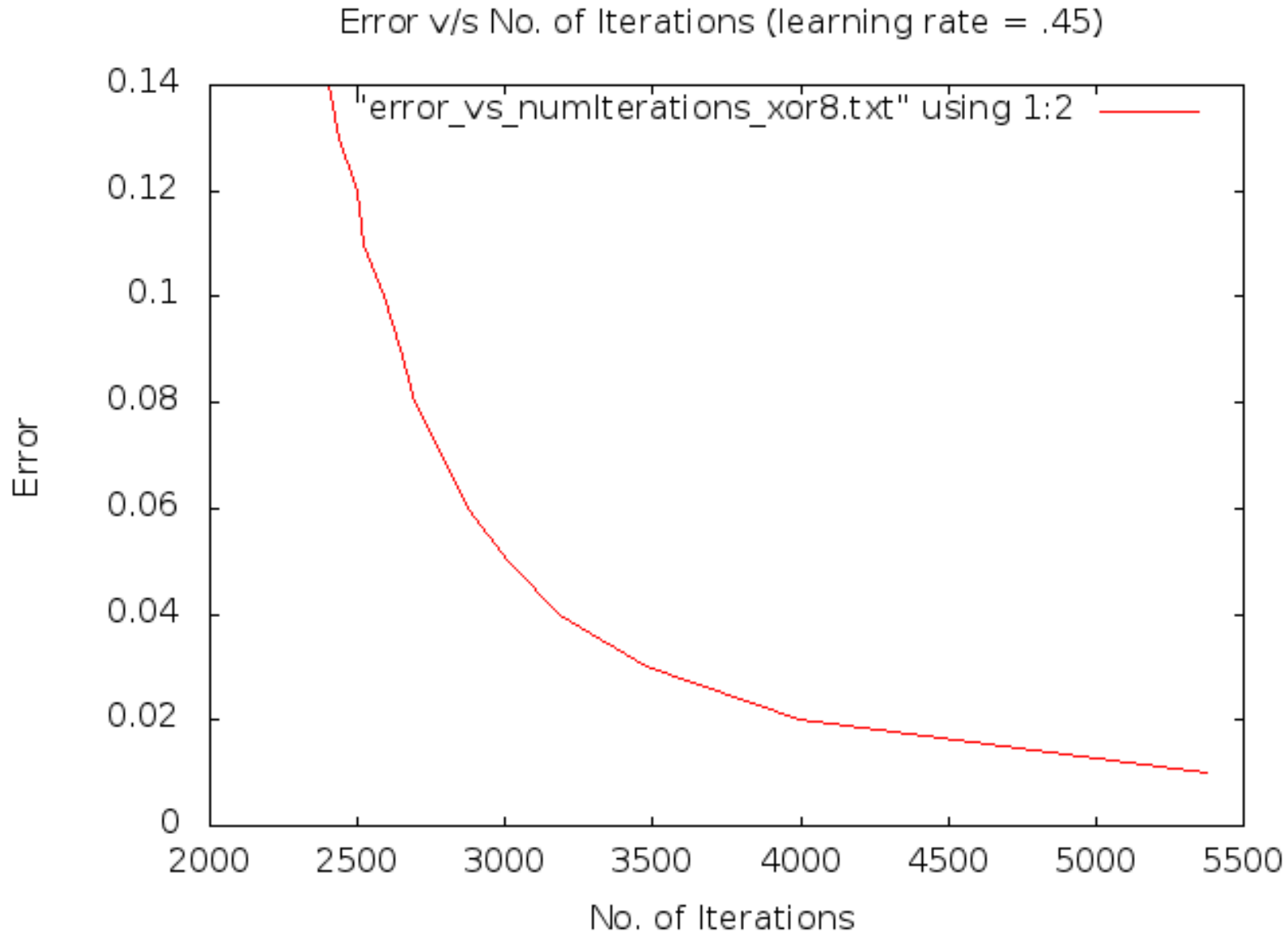
# Effect of Learning Rate (XOR)



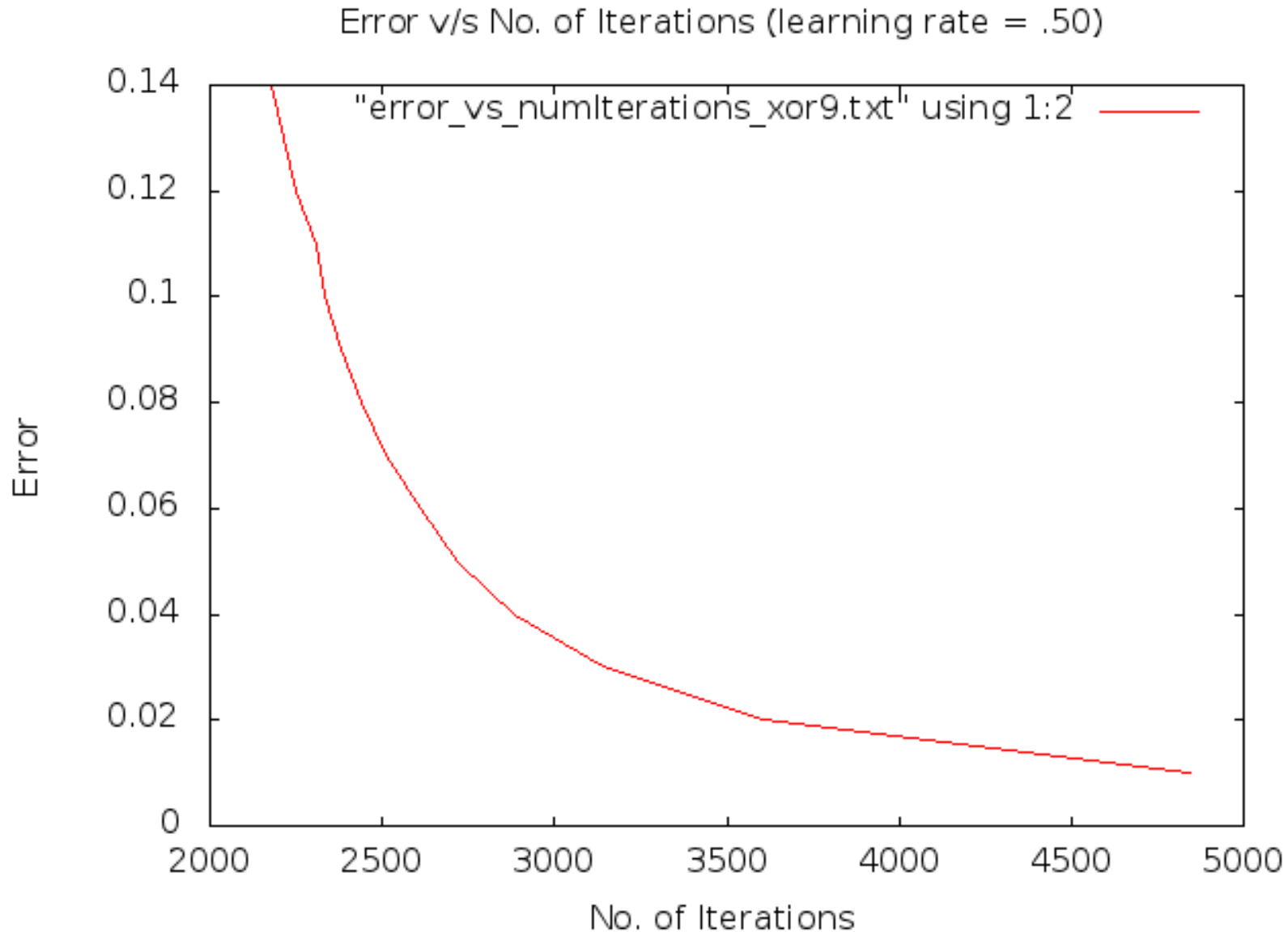
# Effect of Learning Rate (XOR)



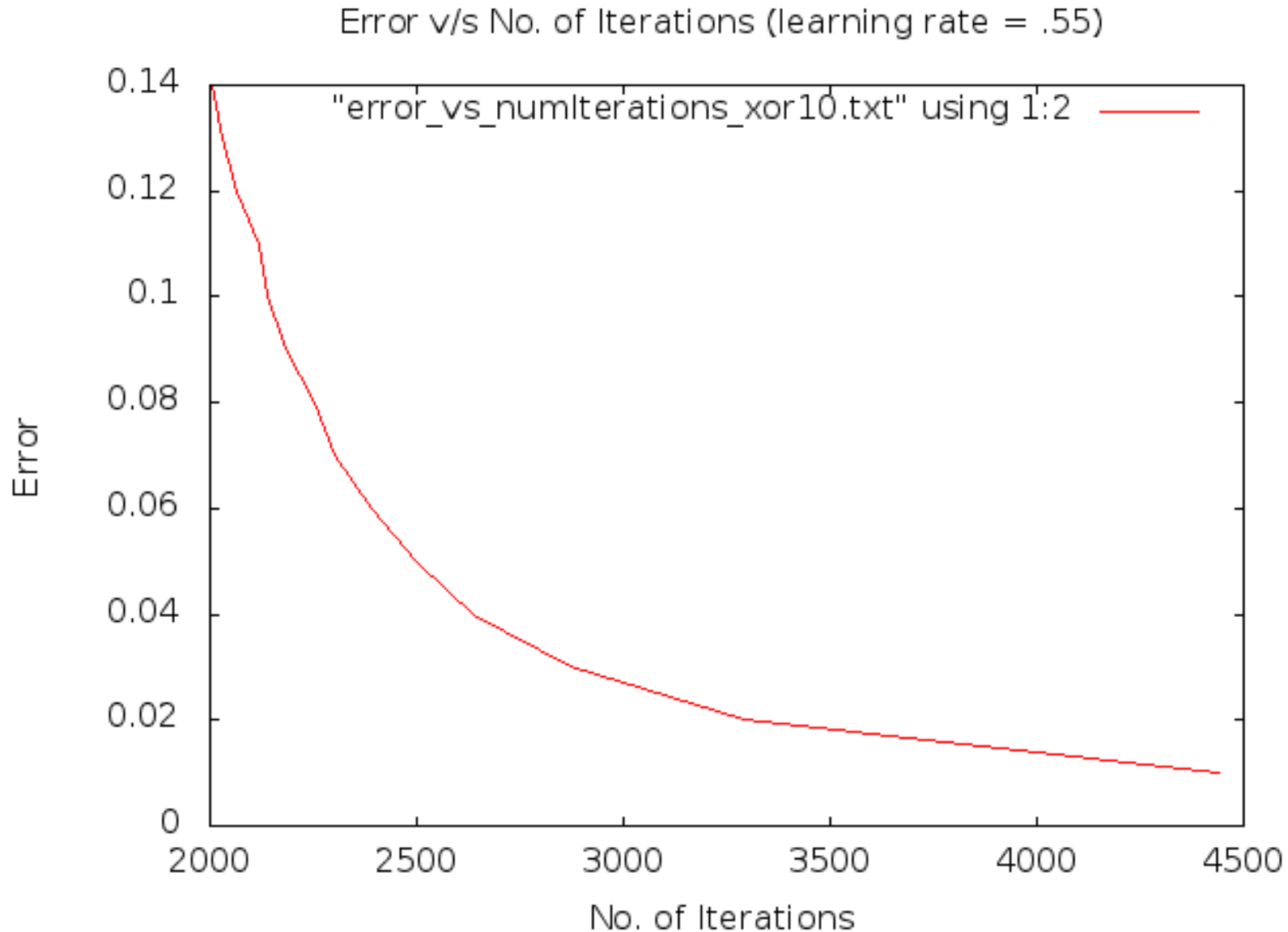
# Effect of Learning Rate (XOR)



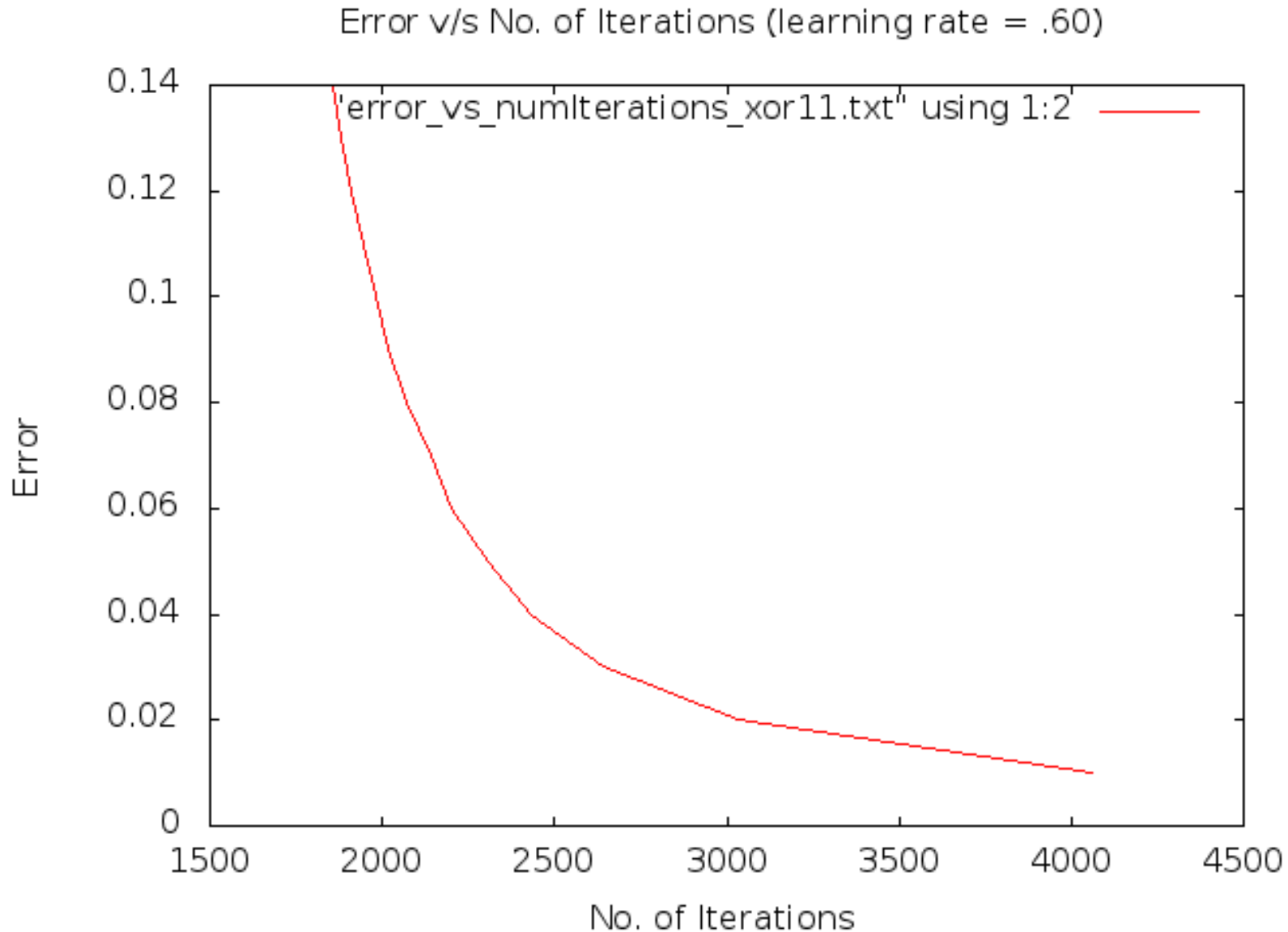
# Effect of Learning Rate (XOR)



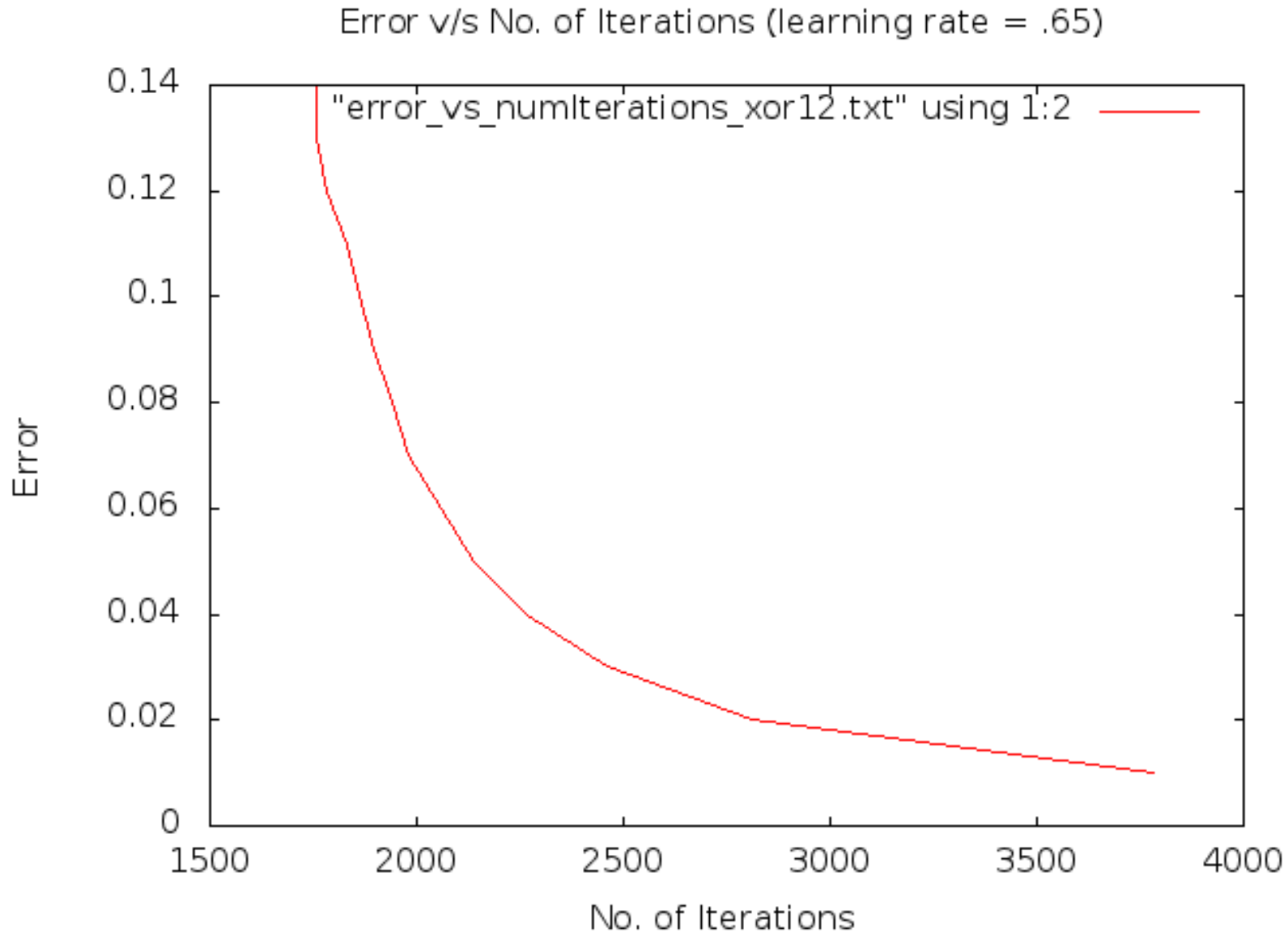
# Effect of Learning Rate (XOR)



# Effect of Learning Rate (XOR)

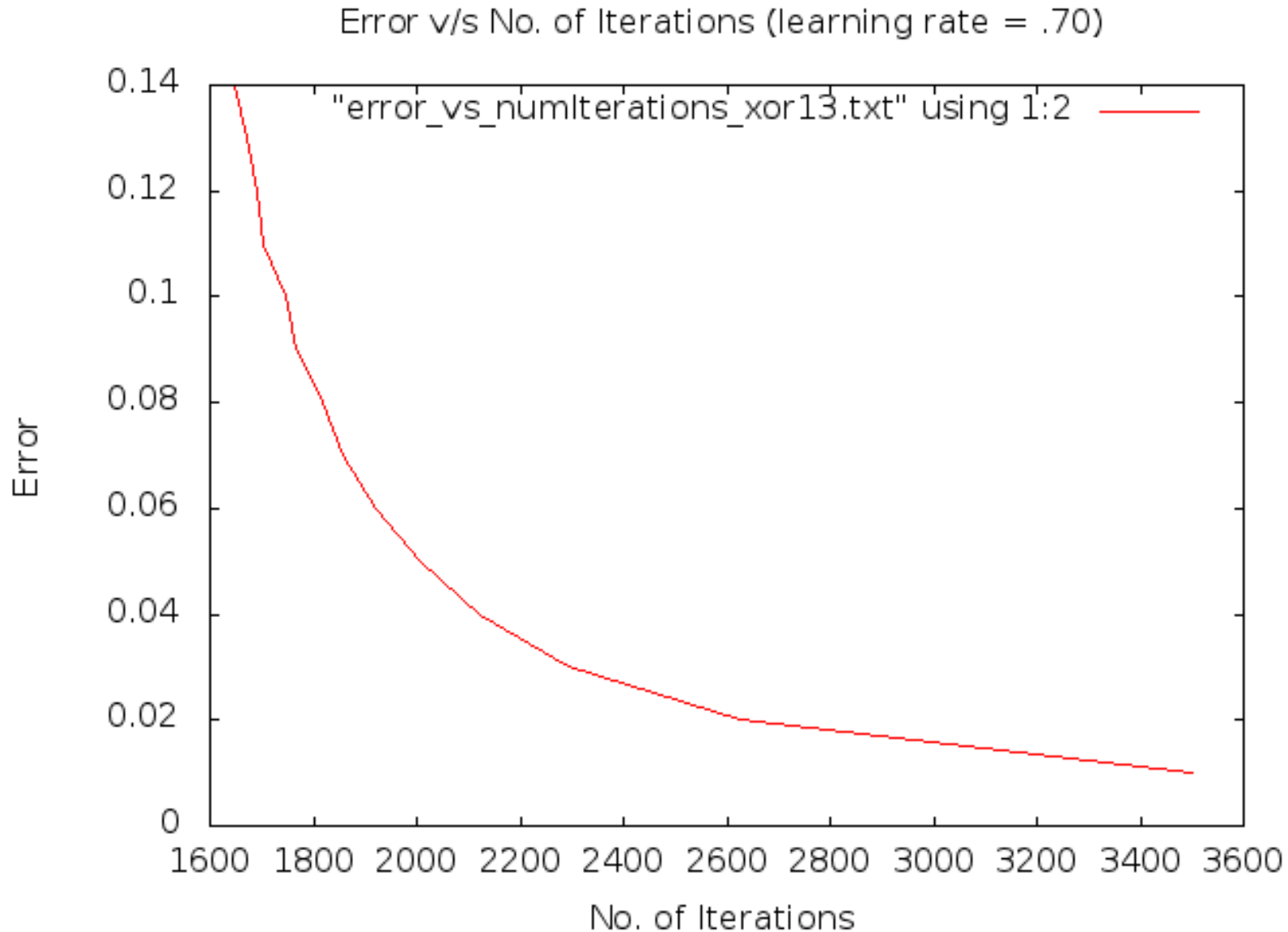


# Effect of Learning Rate (XOR)

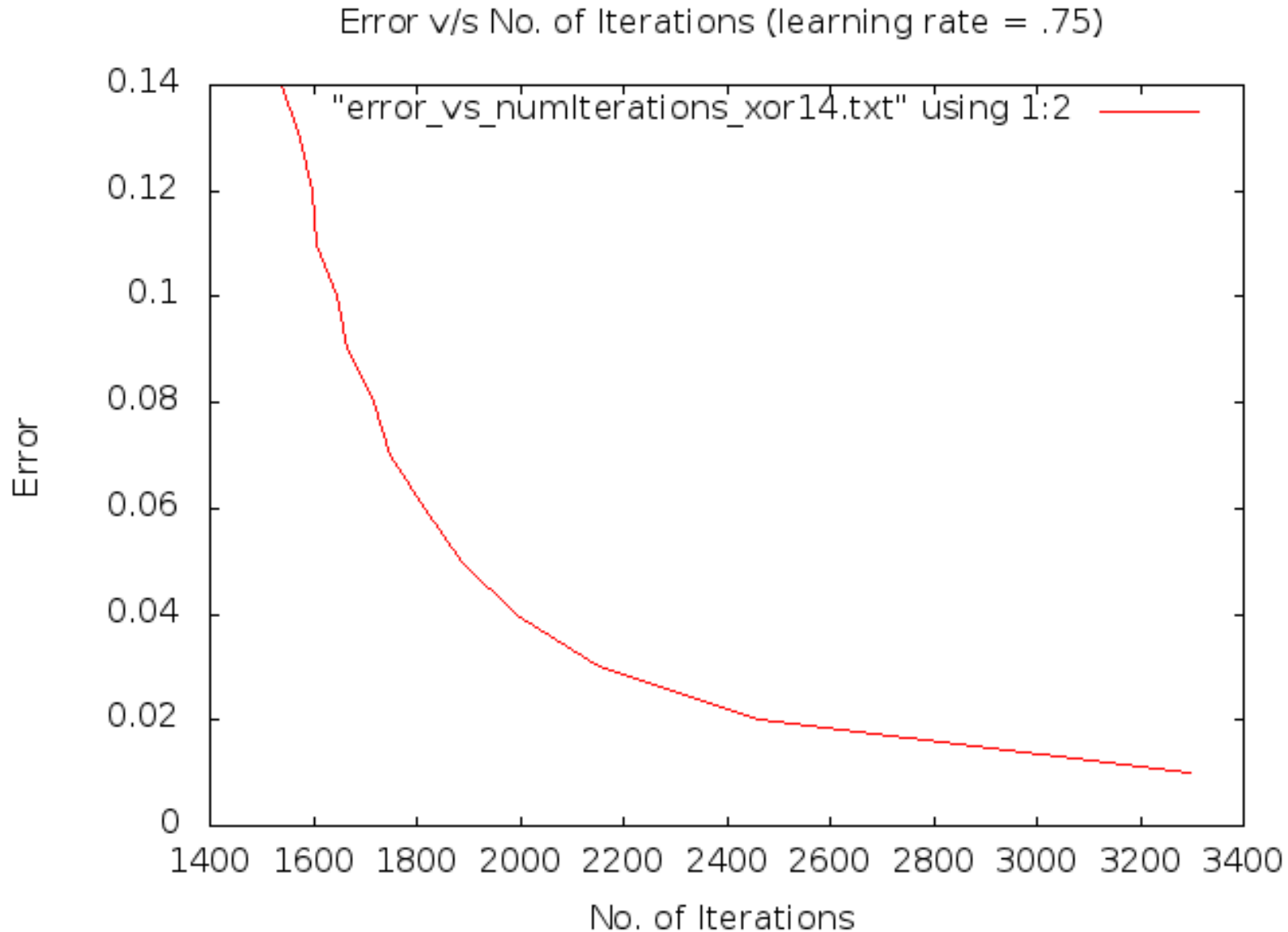




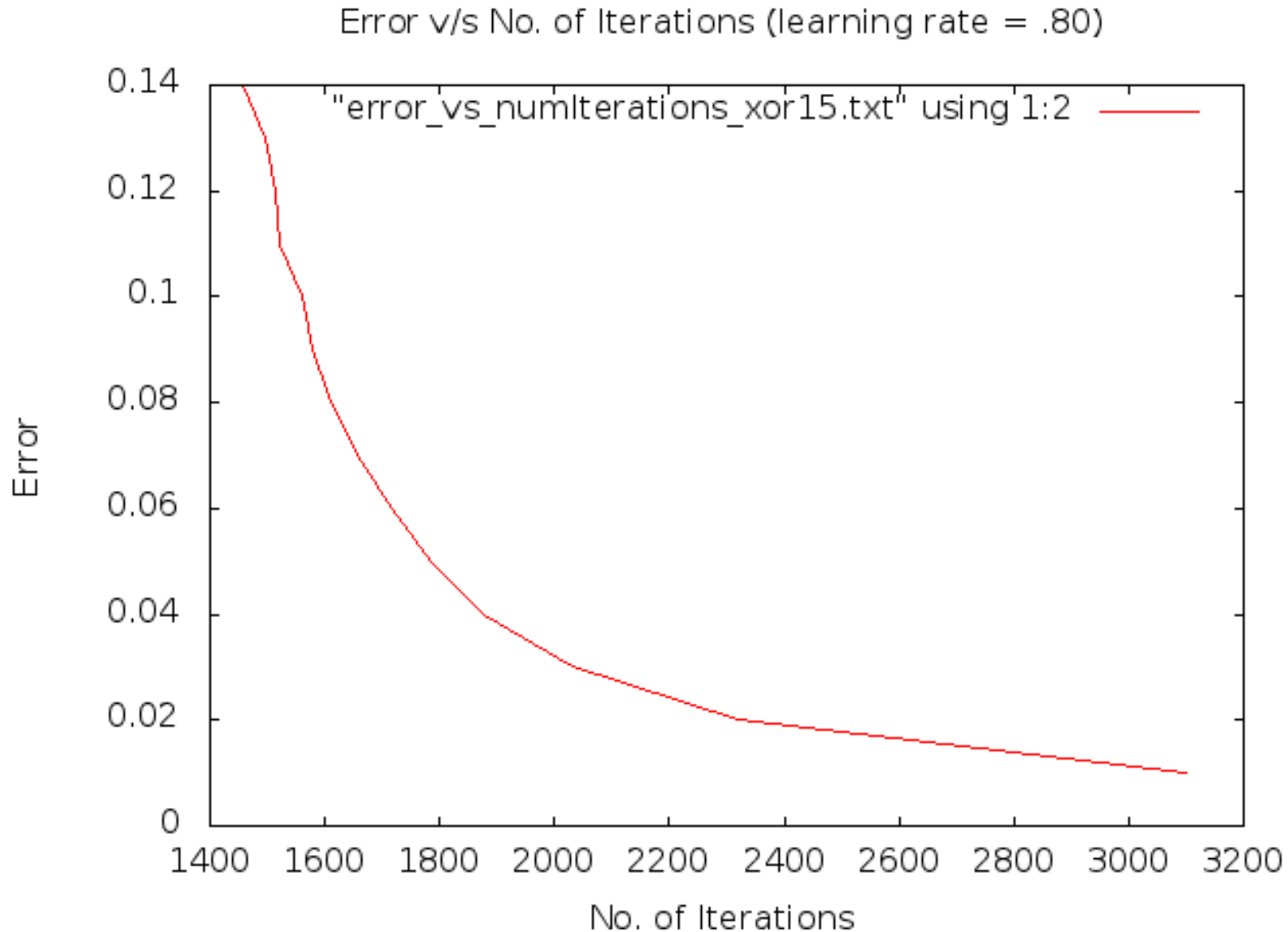
# Effect of Learning Rate (XOR)



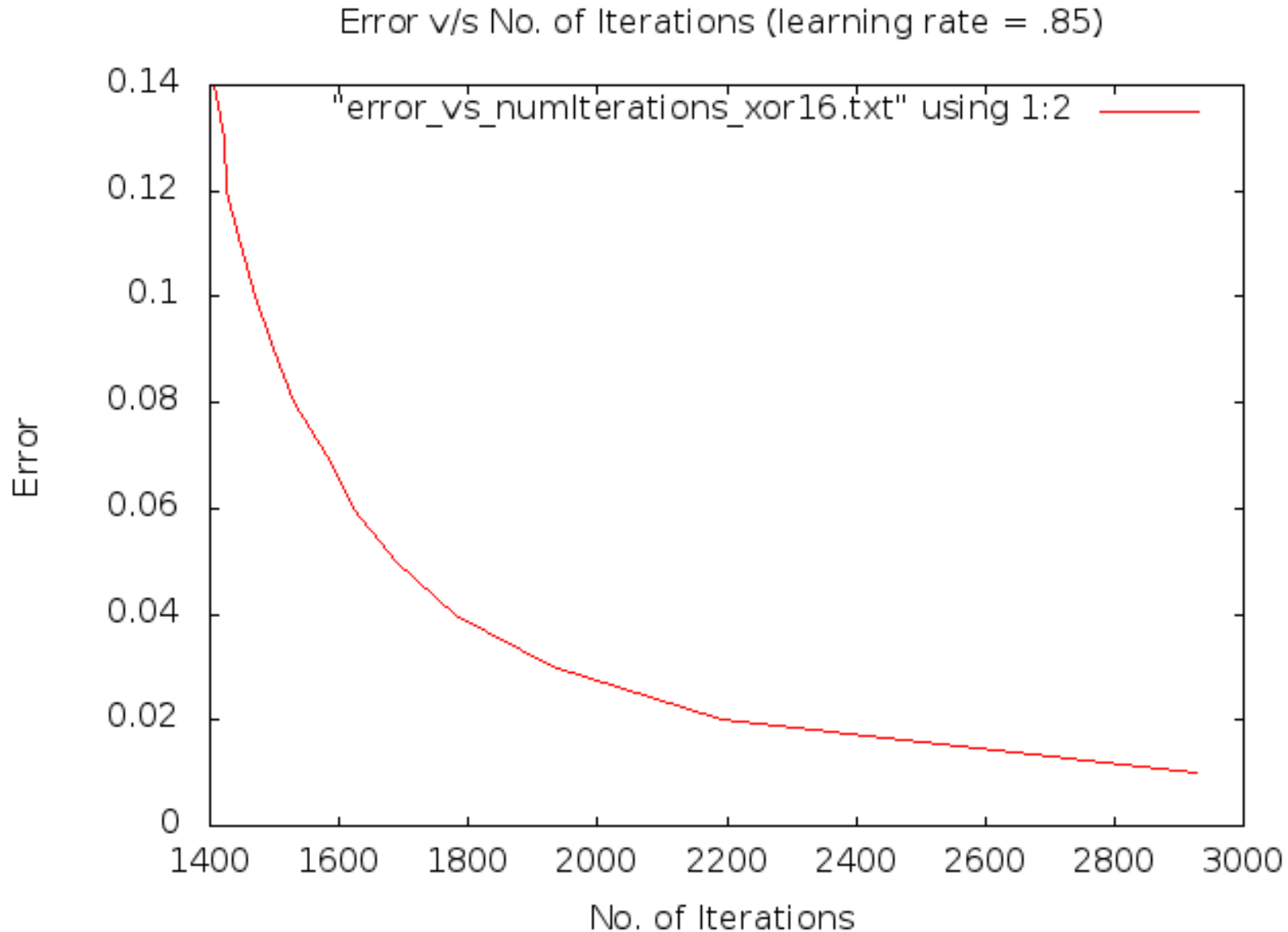
# Effect of Learning Rate (XOR)



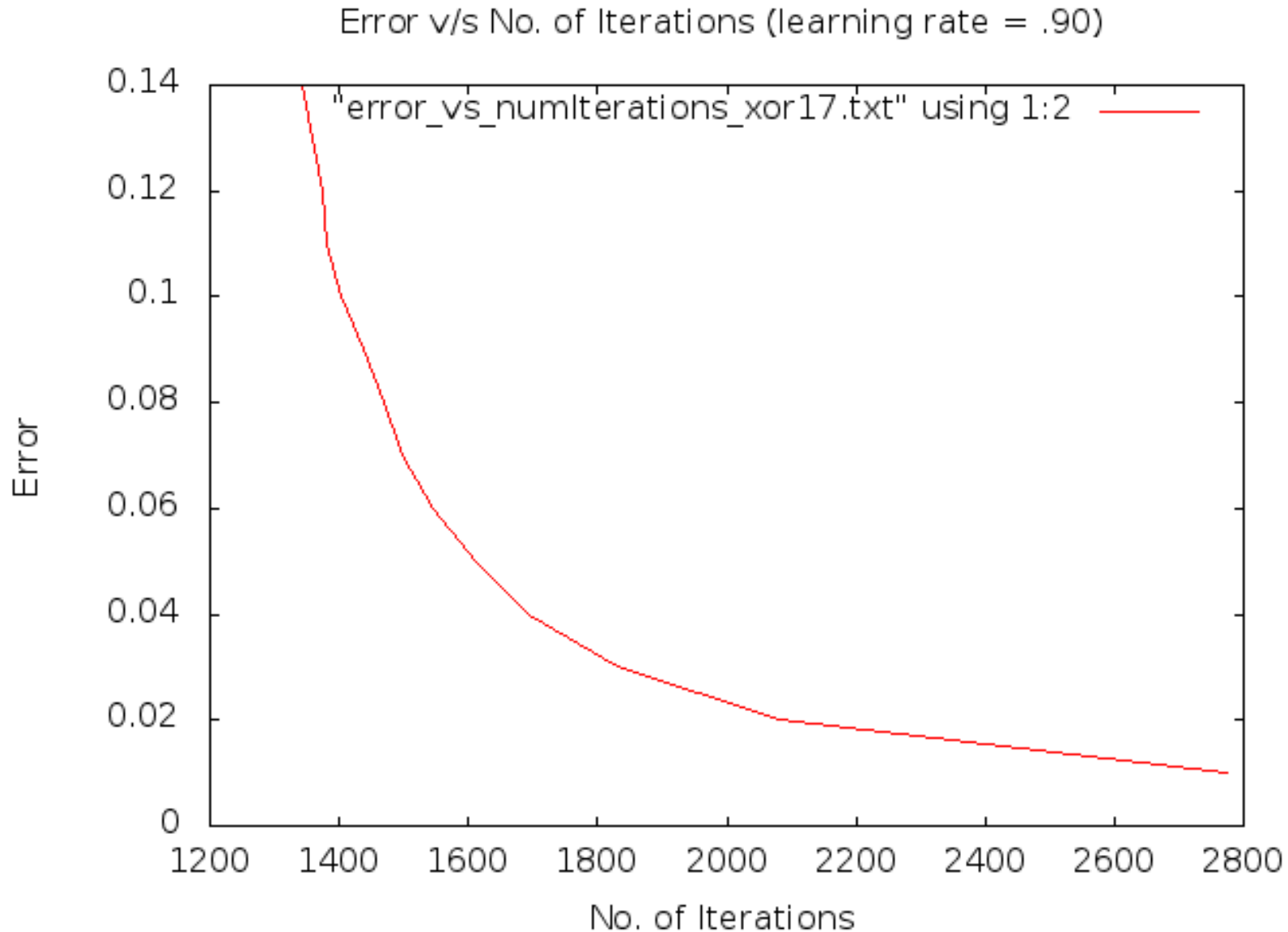
# Effect of Learning Rate (XOR)



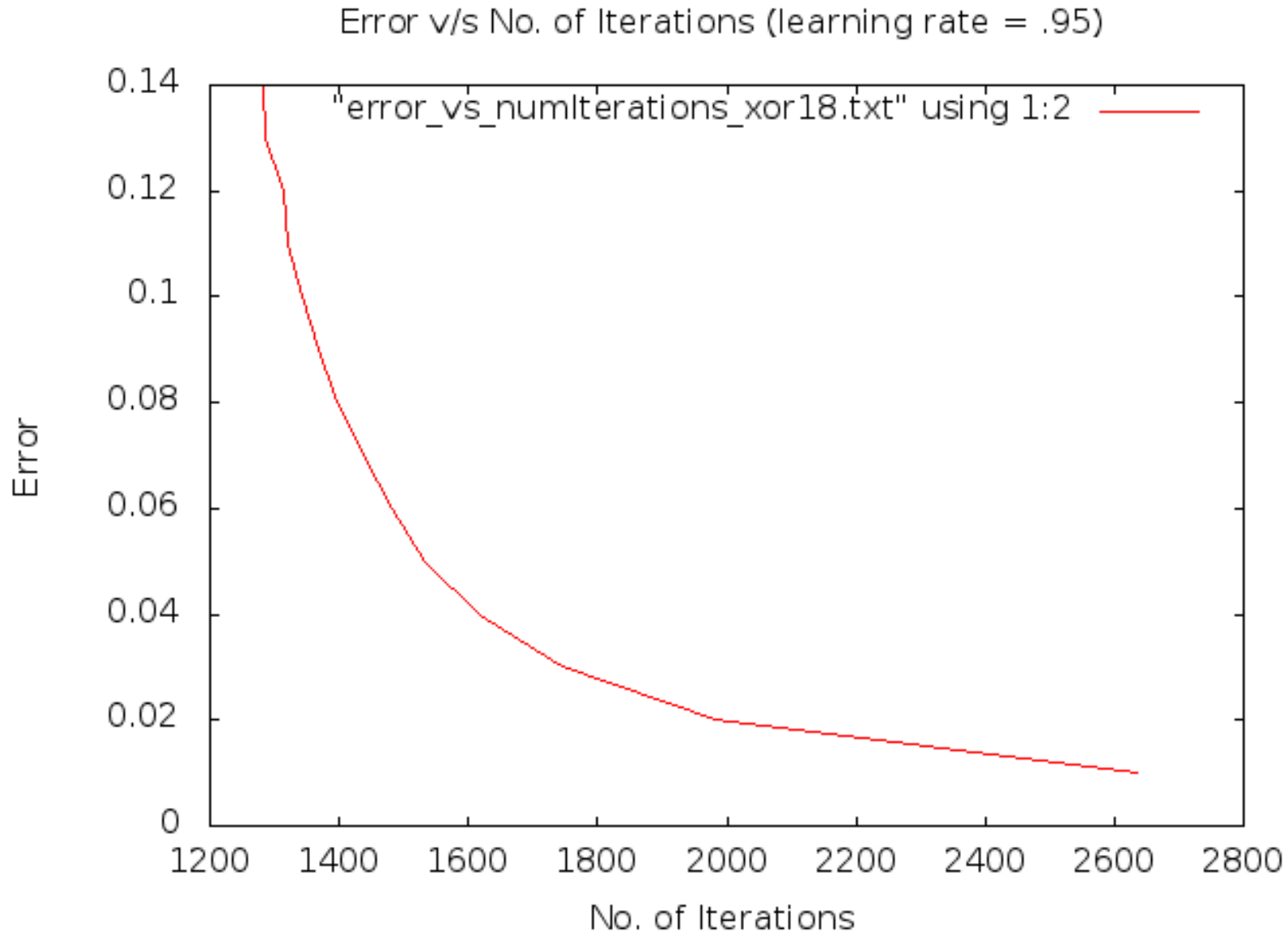
# Effect of Learning Rate (XOR)



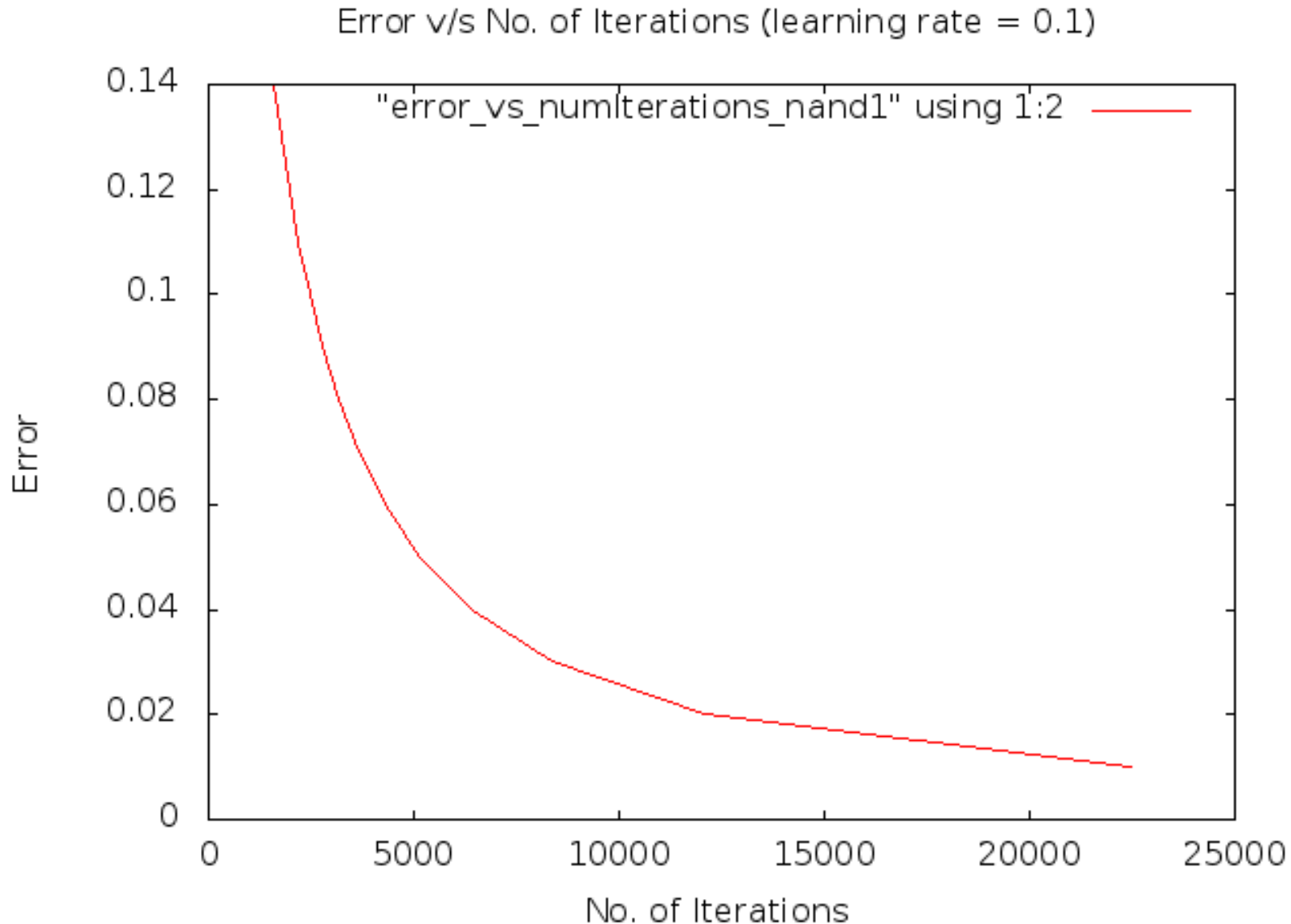
# Effect of Learning Rate (XOR)



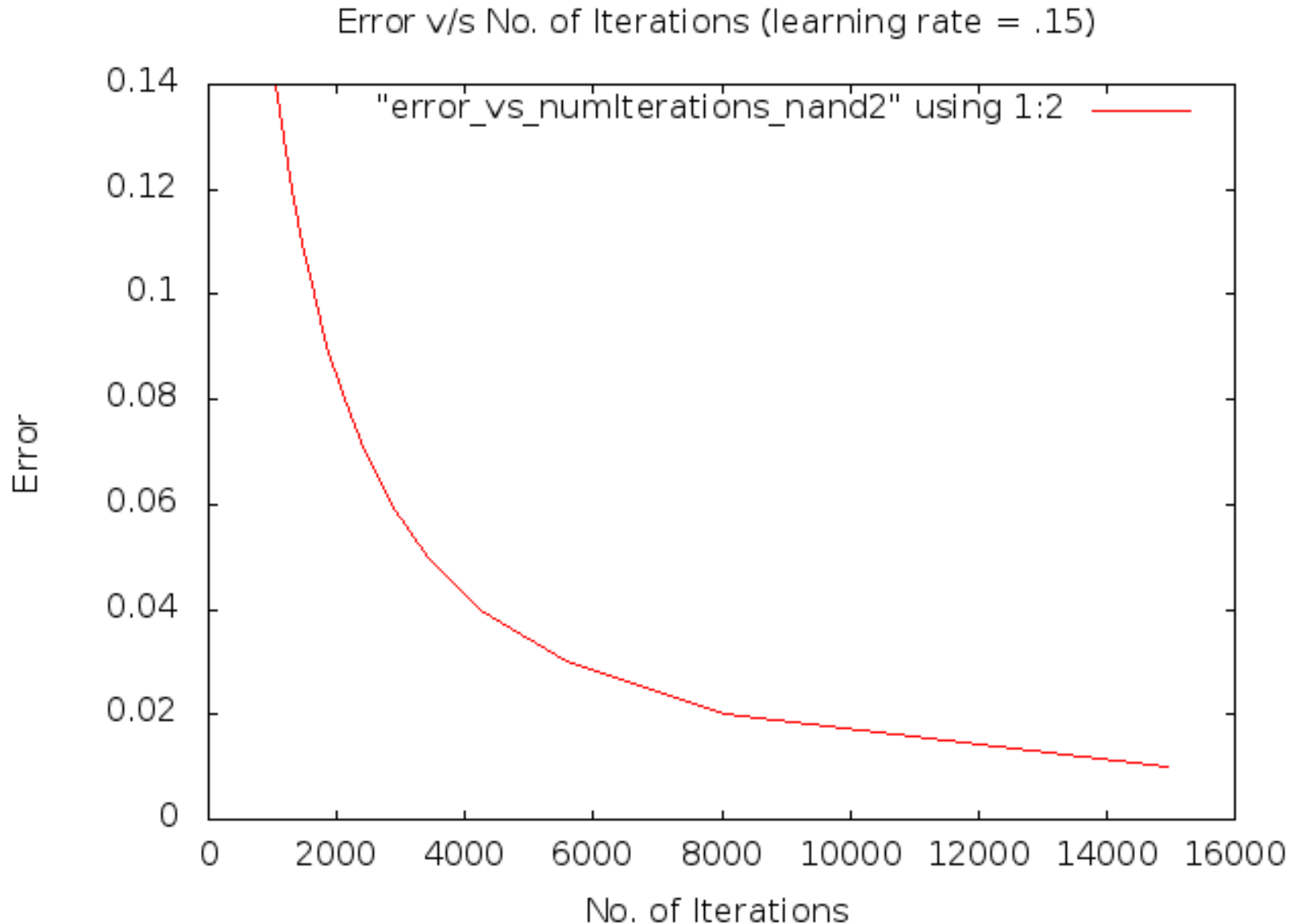
# Effect of Learning Rate (XOR)



# Effect of Learning Rate (NAND)

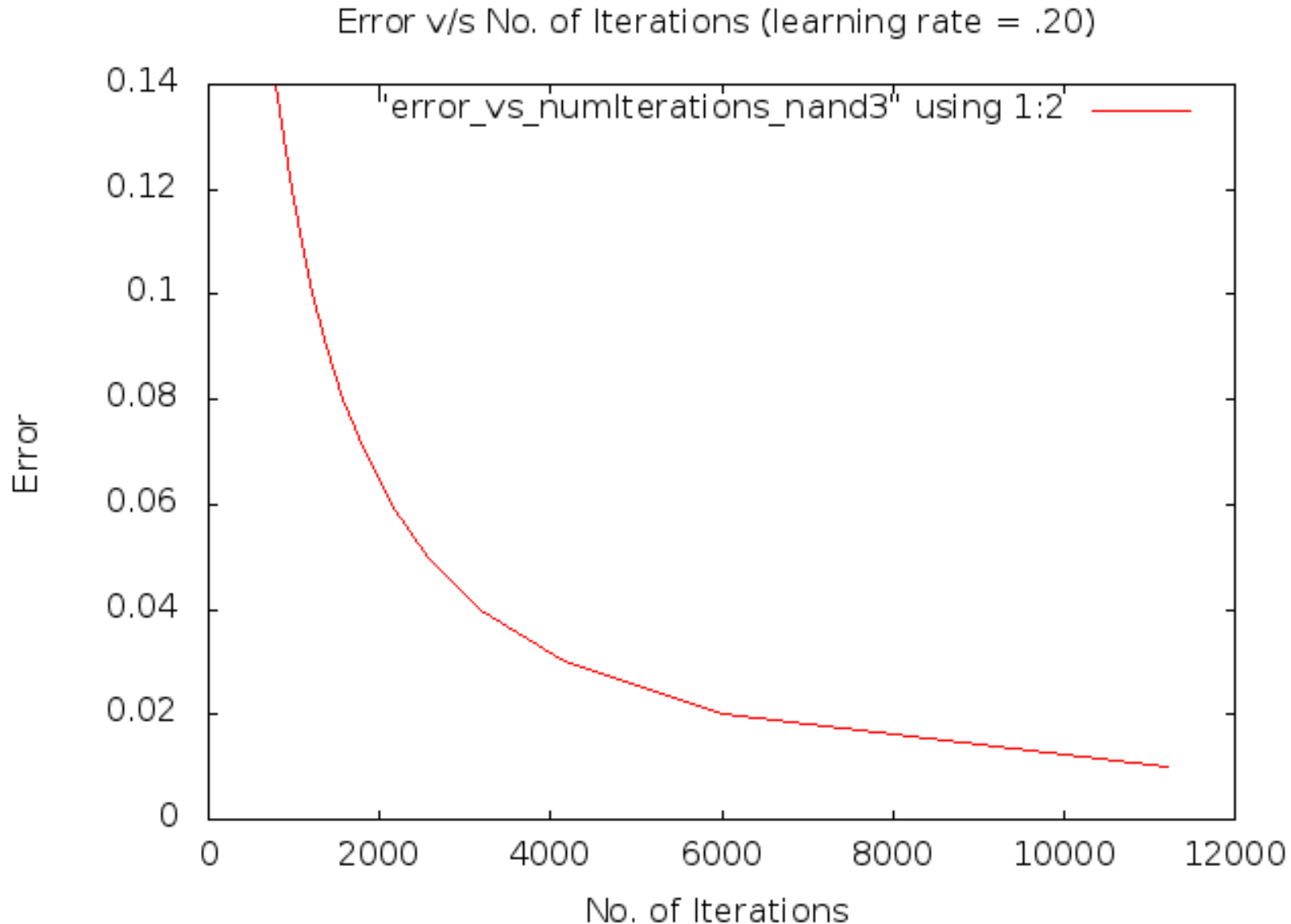


# Effect of Learning Rate (NAND)

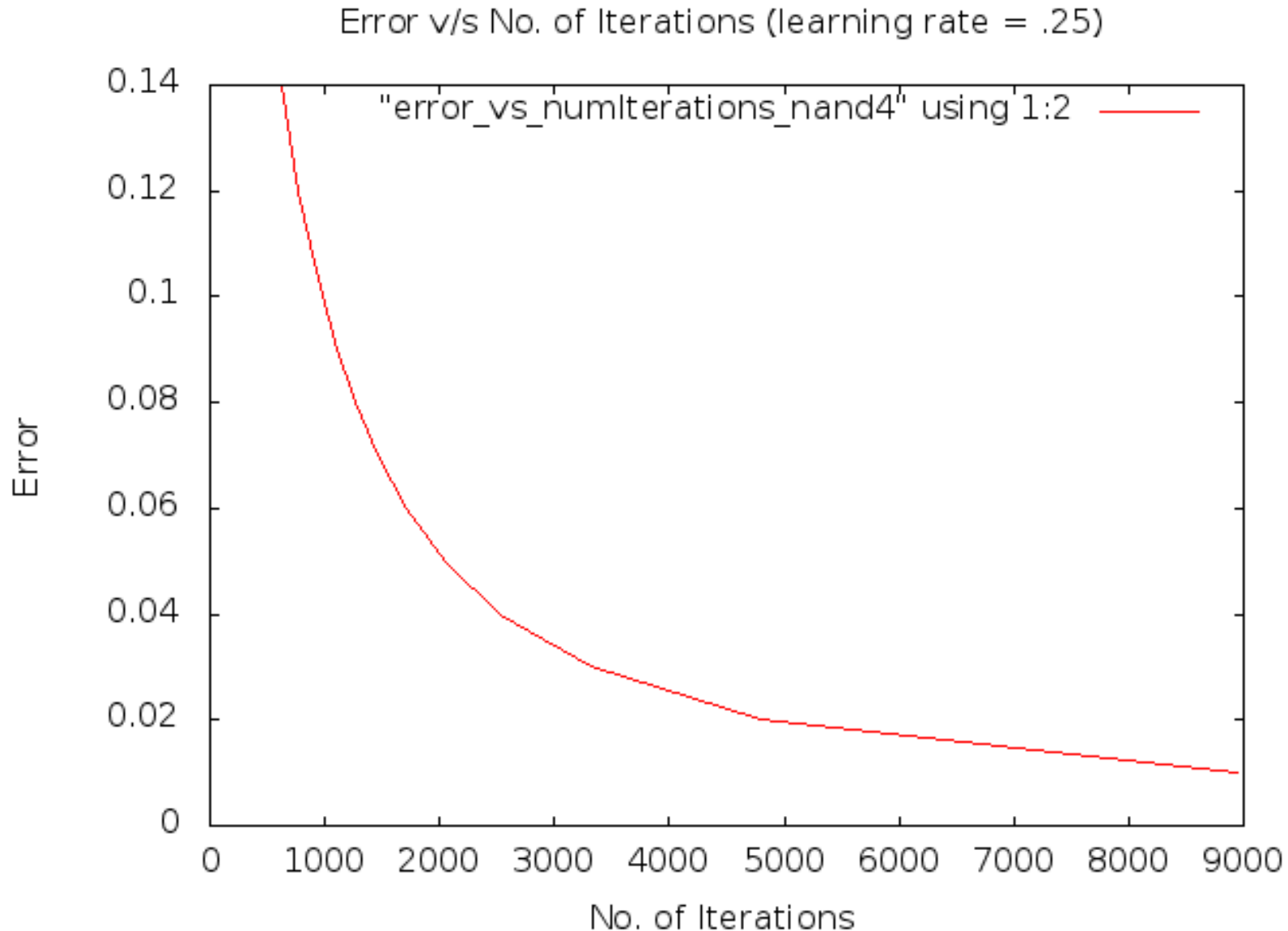




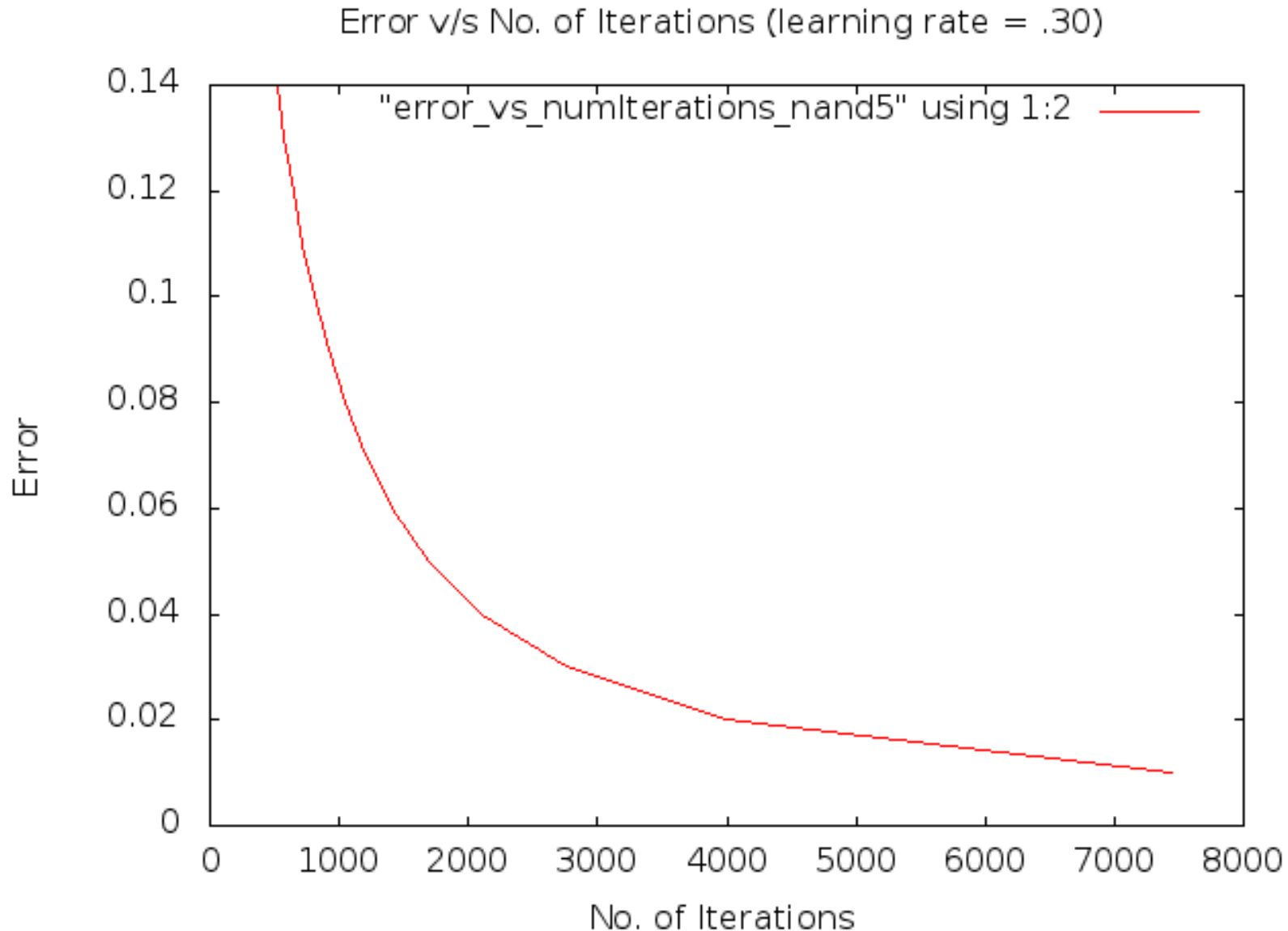
# Effect of Learning Rate (NAND)



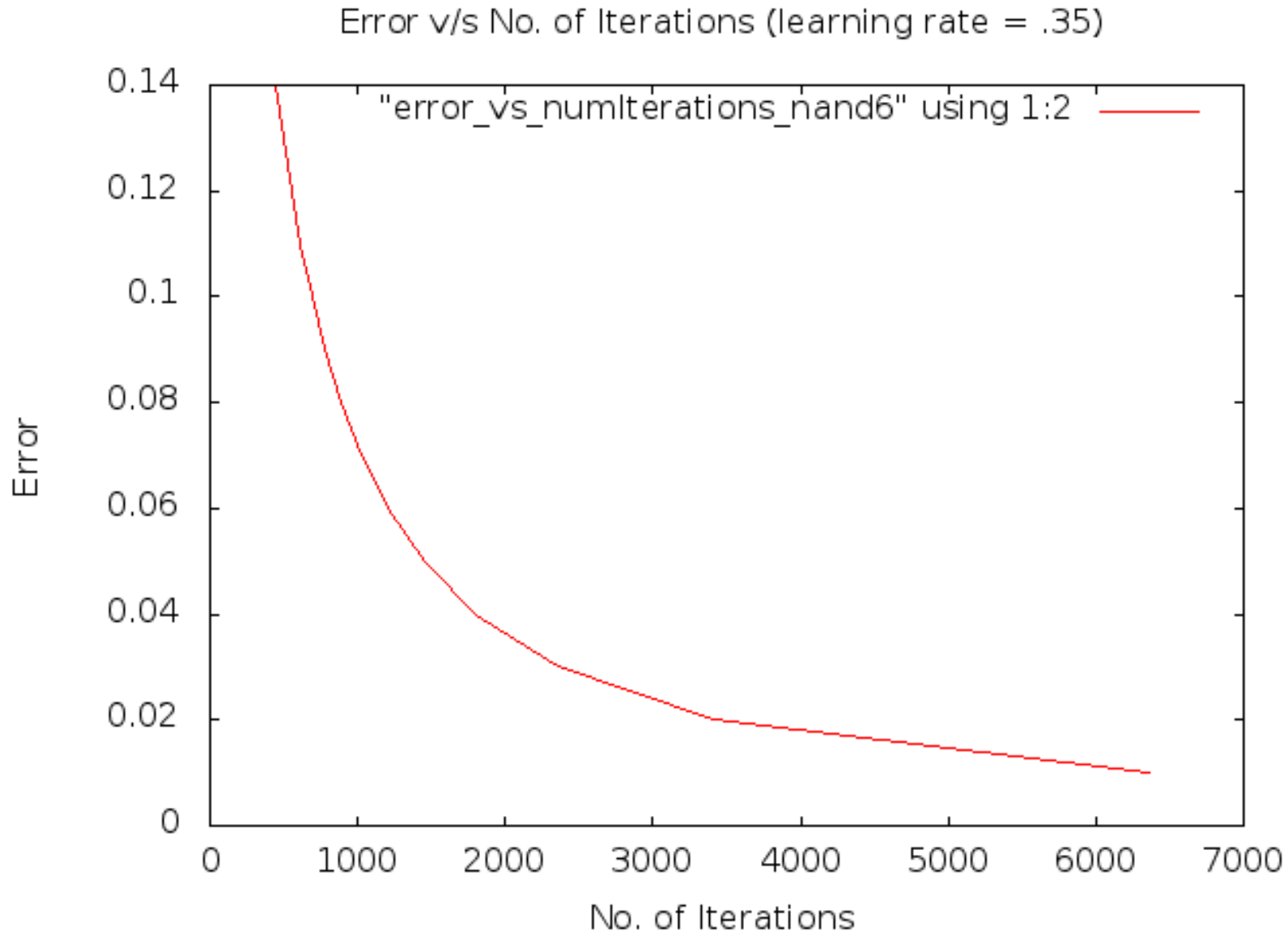
# Effect of Learning Rate (NAND)



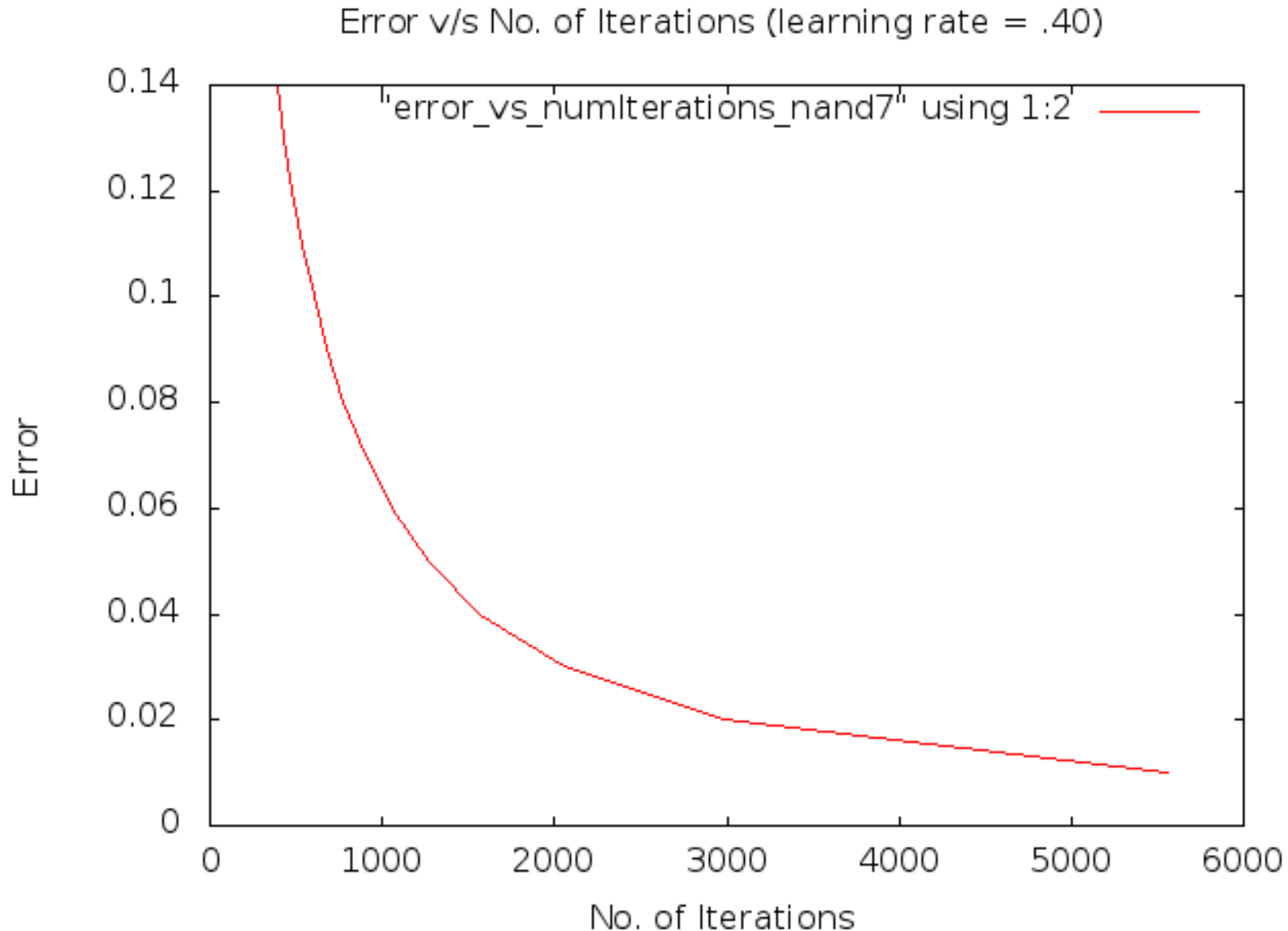
# Effect of Learning Rate (NAND)



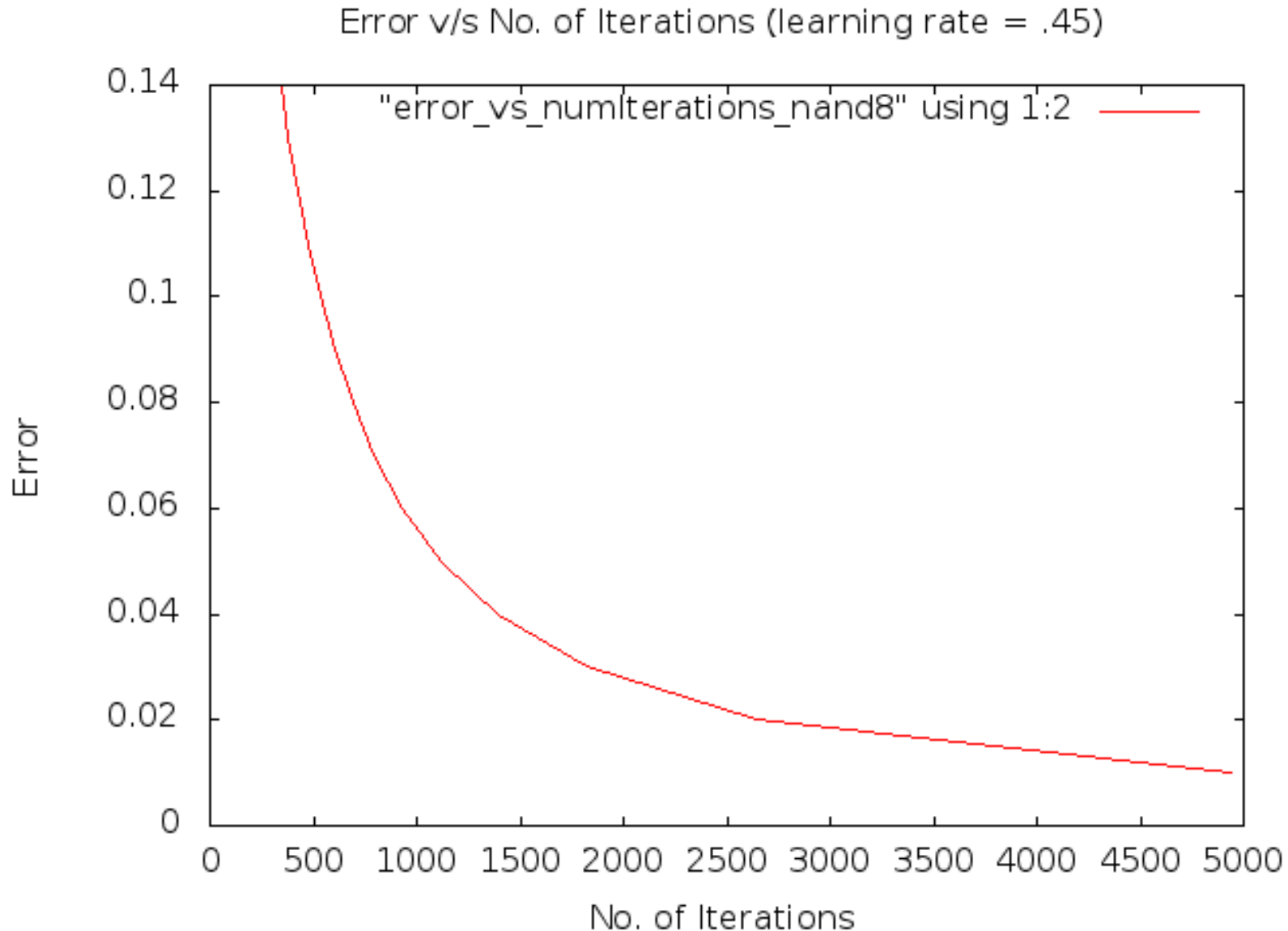
# Effect of Learning Rate (NAND)



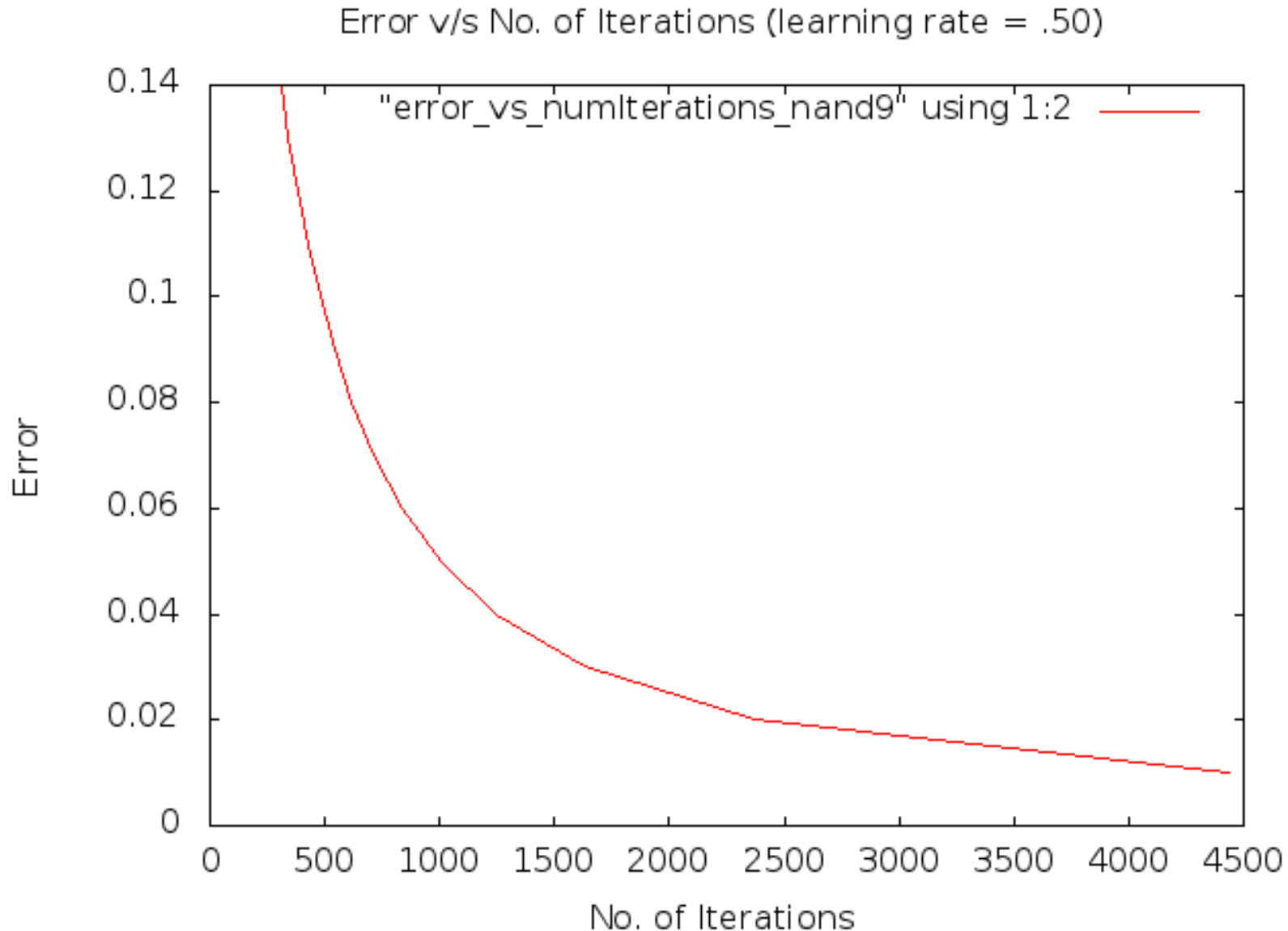
# Effect of Learning Rate (NAND)



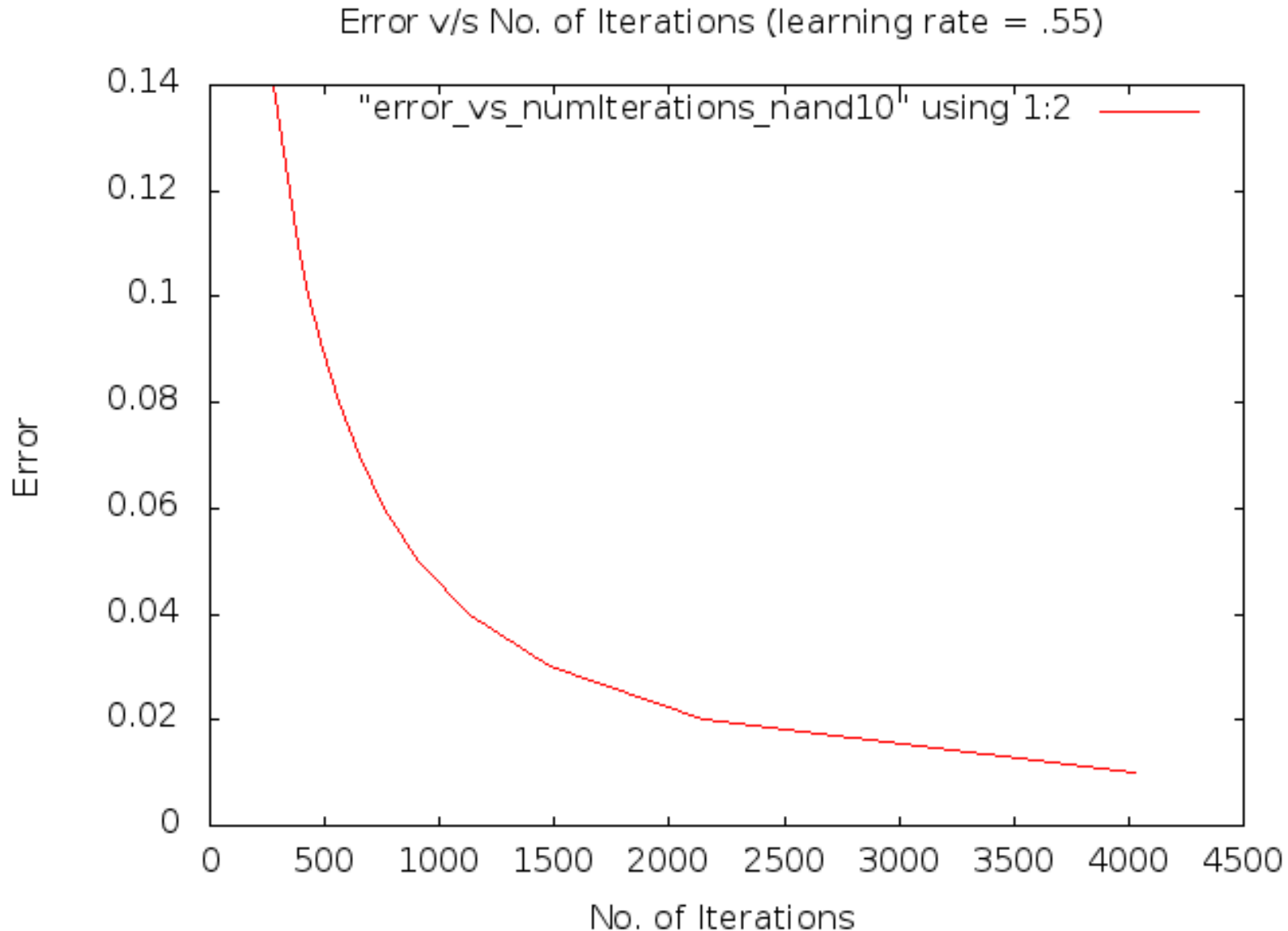
# Effect of Learning Rate (NAND)



# Effect of Learning Rate (NAND)

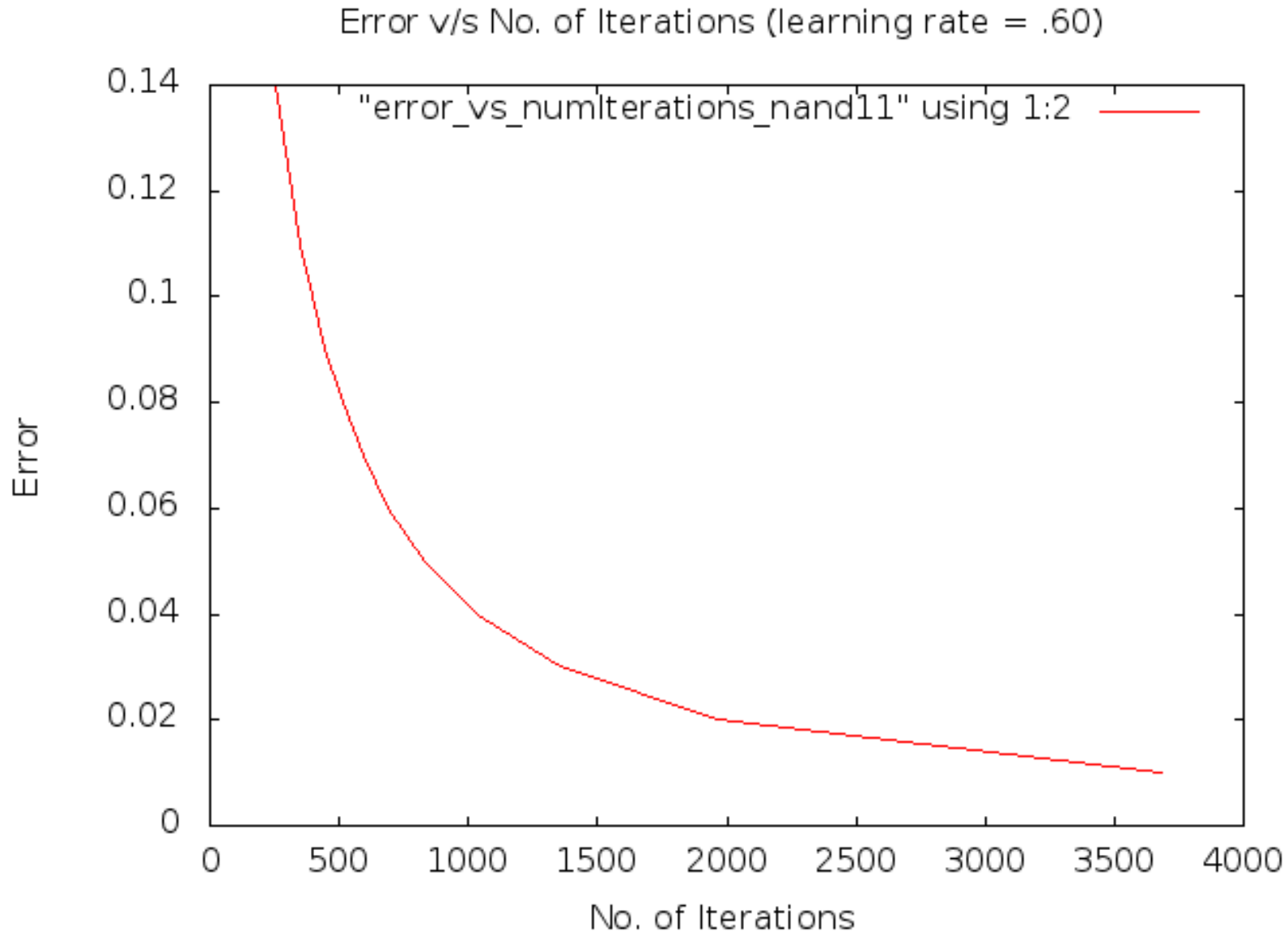


# Effect of Learning Rate (NAND)

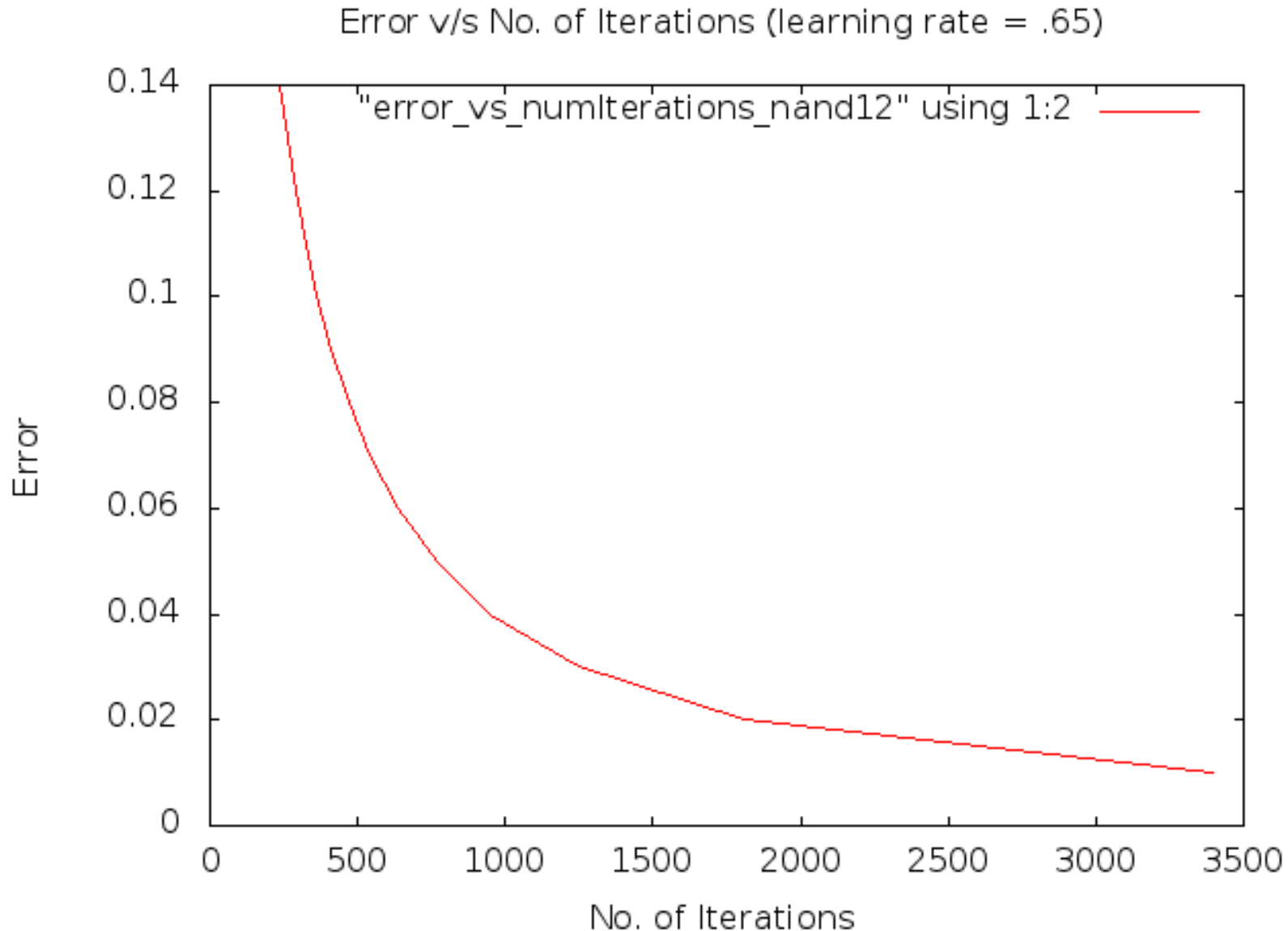




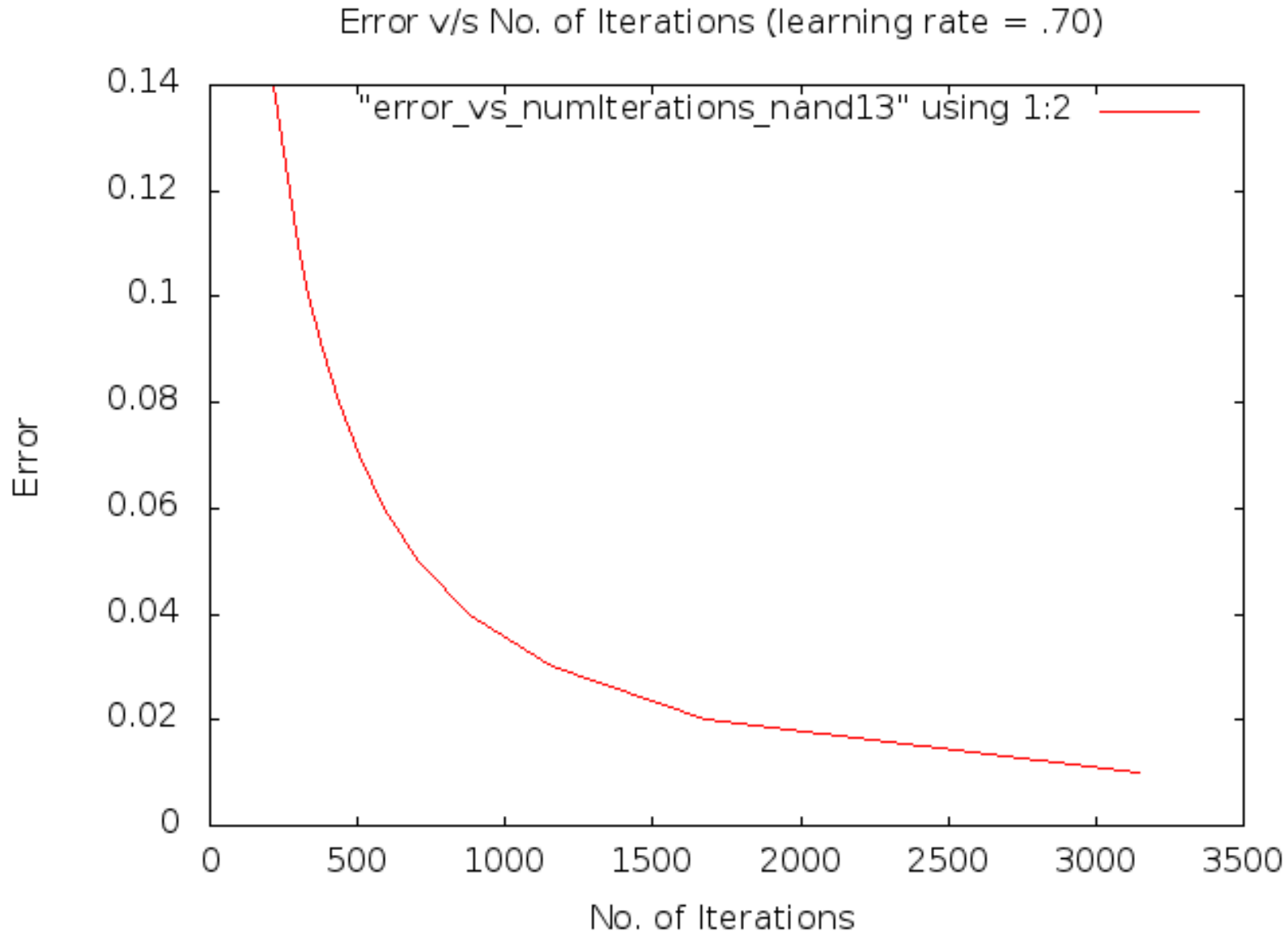
# Effect of Learning Rate (NAND)



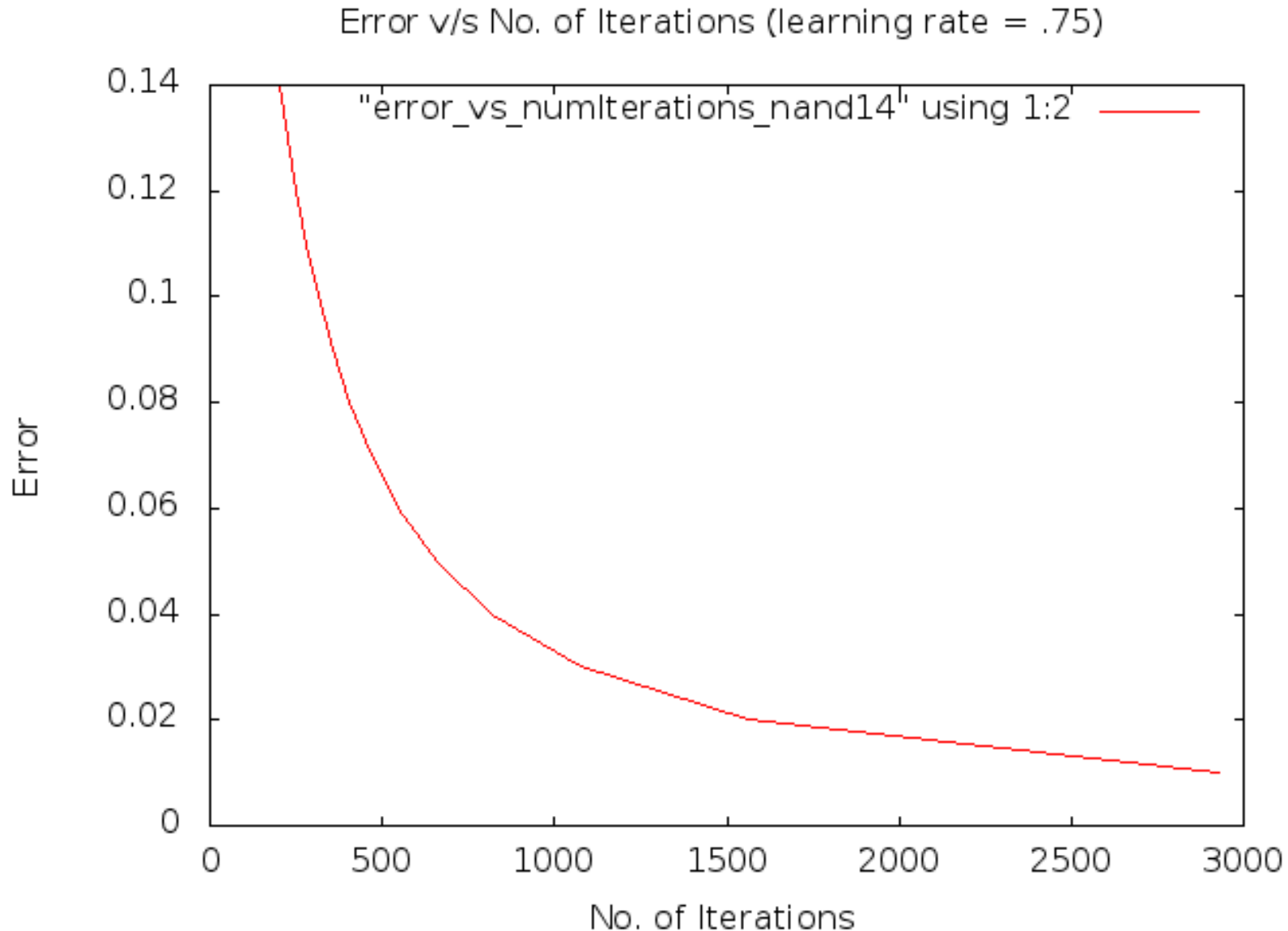
# Effect of Learning Rate (NAND)



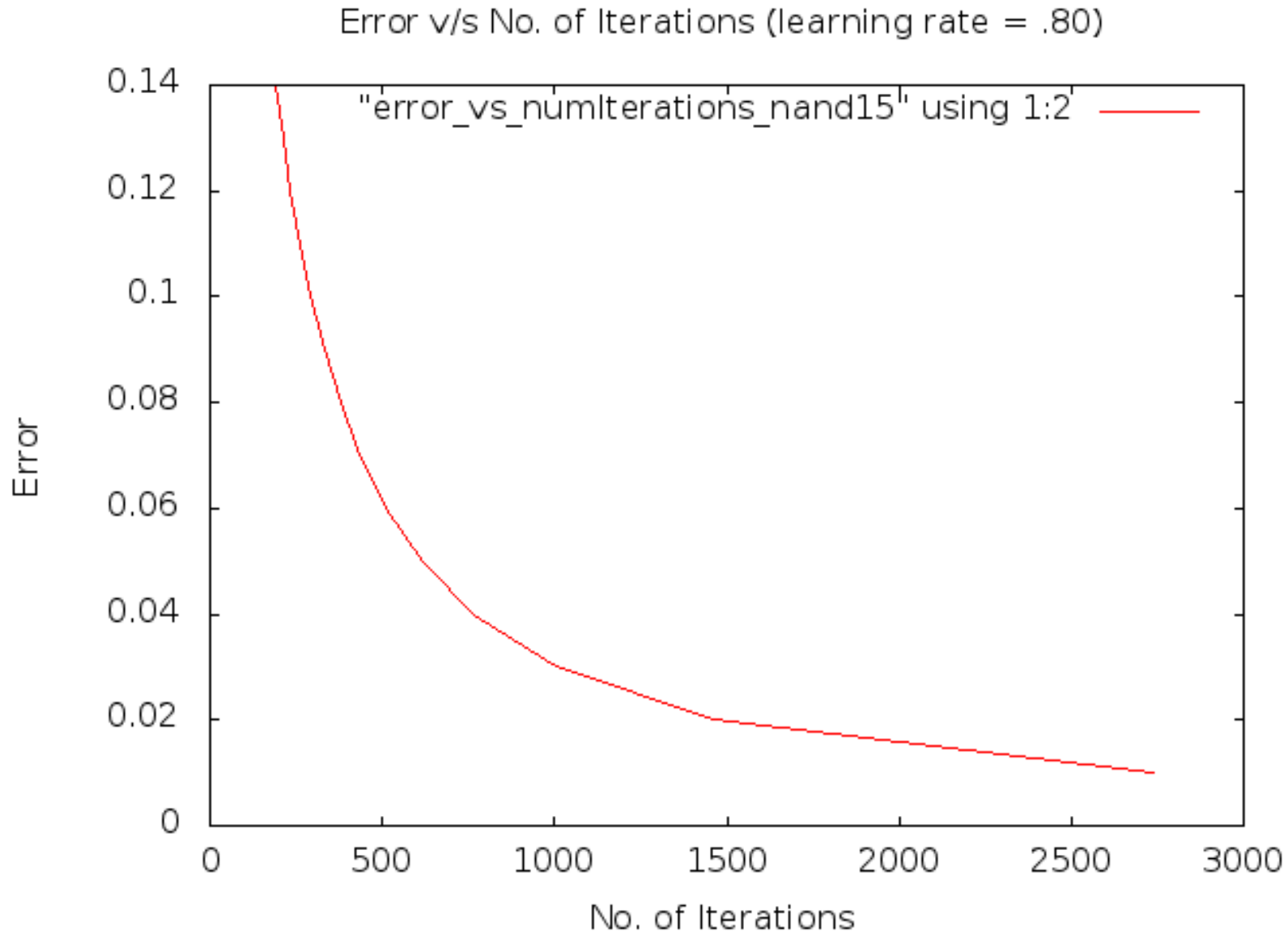
# Effect of Learning Rate (NAND)



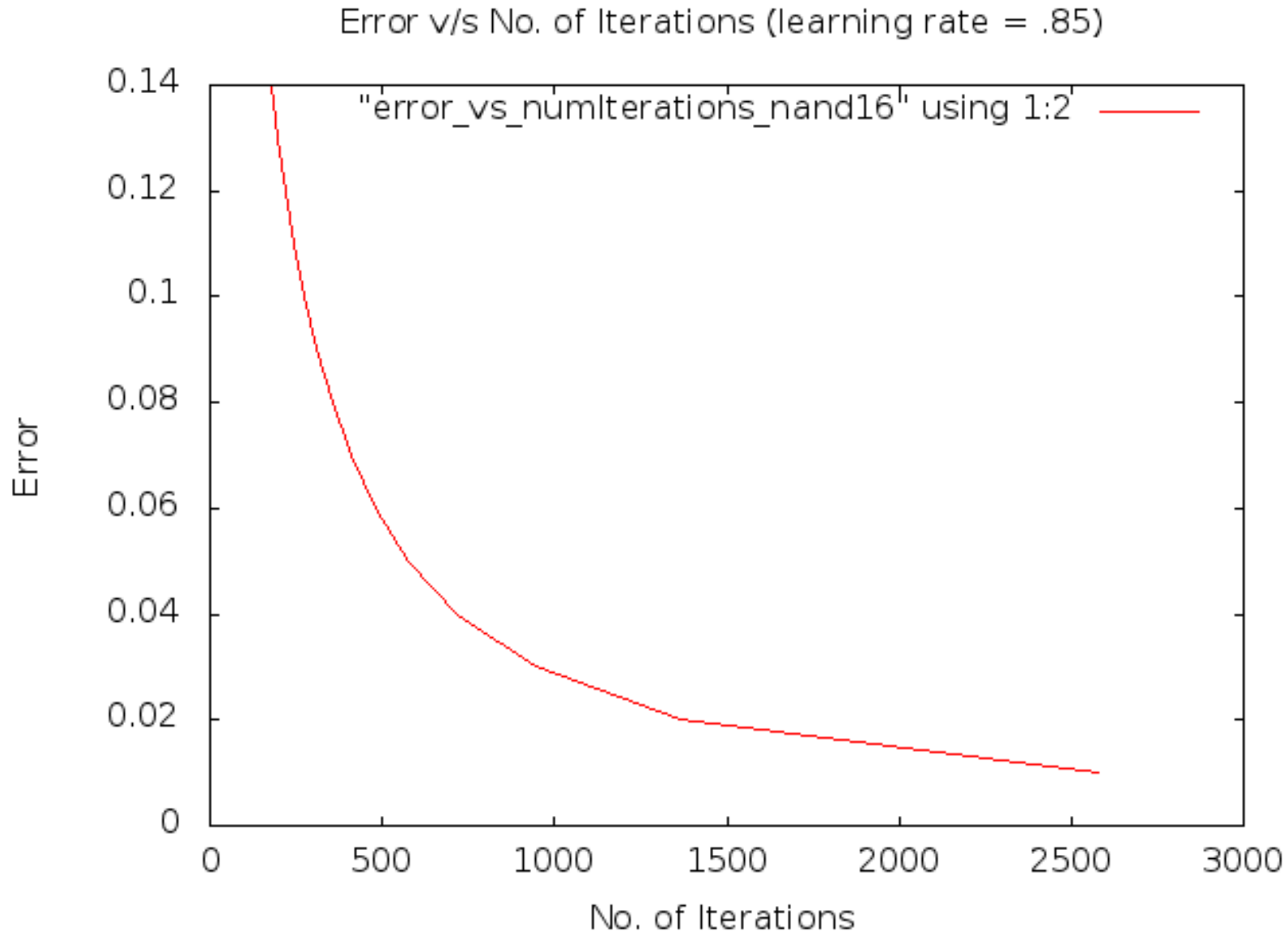
# Effect of Learning Rate (NAND)



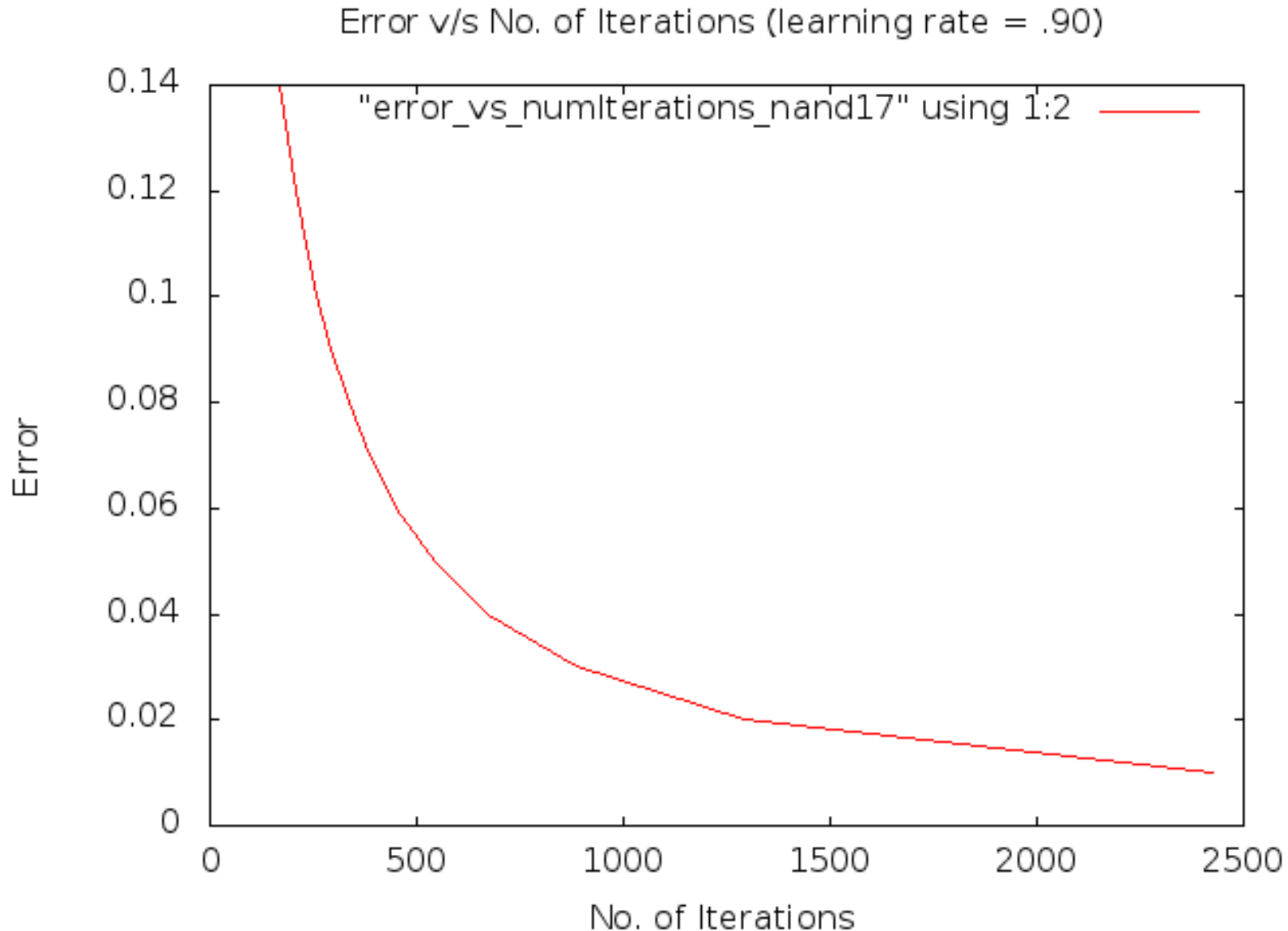
# Effect of Learning Rate (NAND)



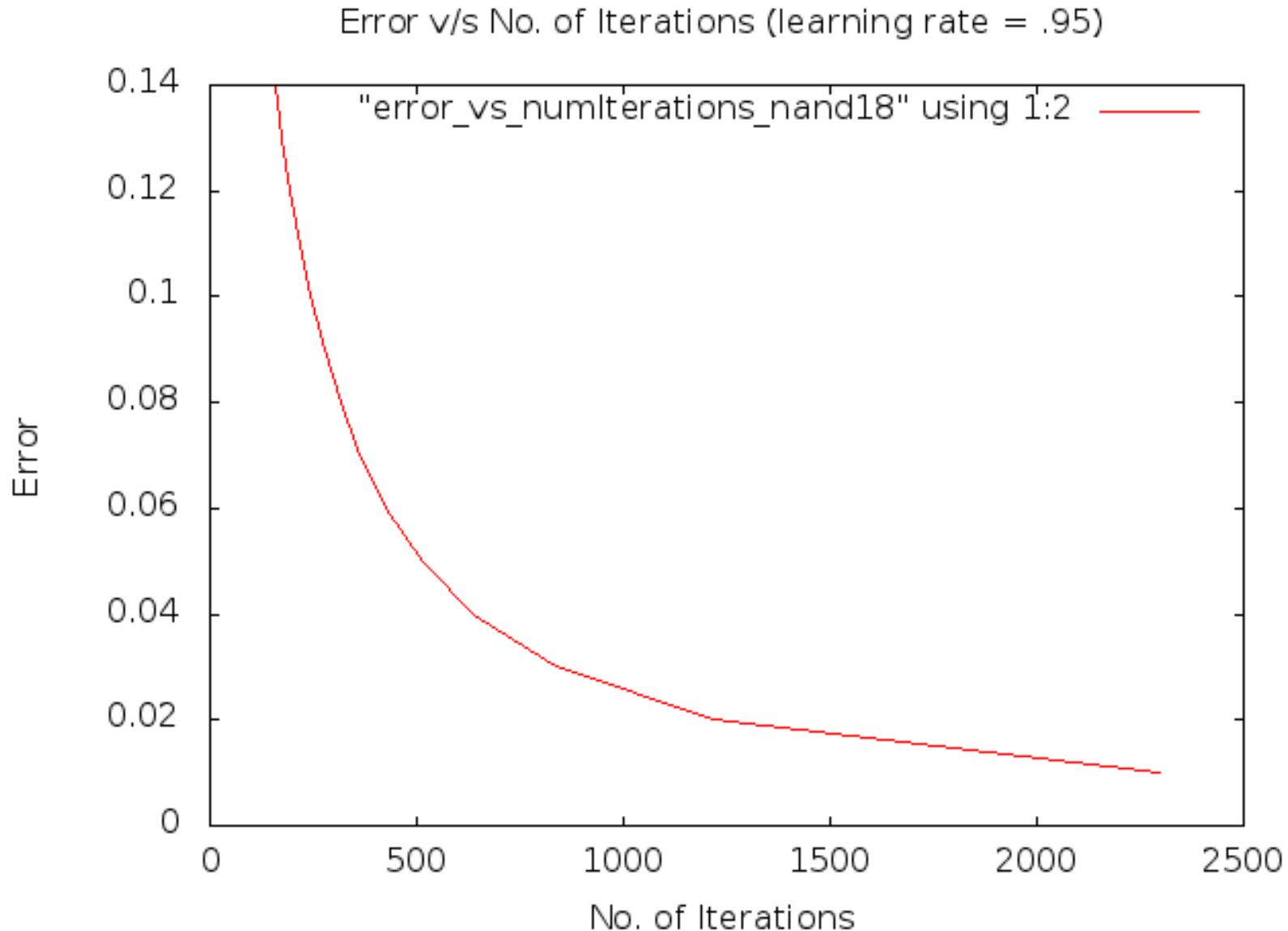
# Effect of Learning Rate (NAND)



# Effect of Learning Rate (NAND)

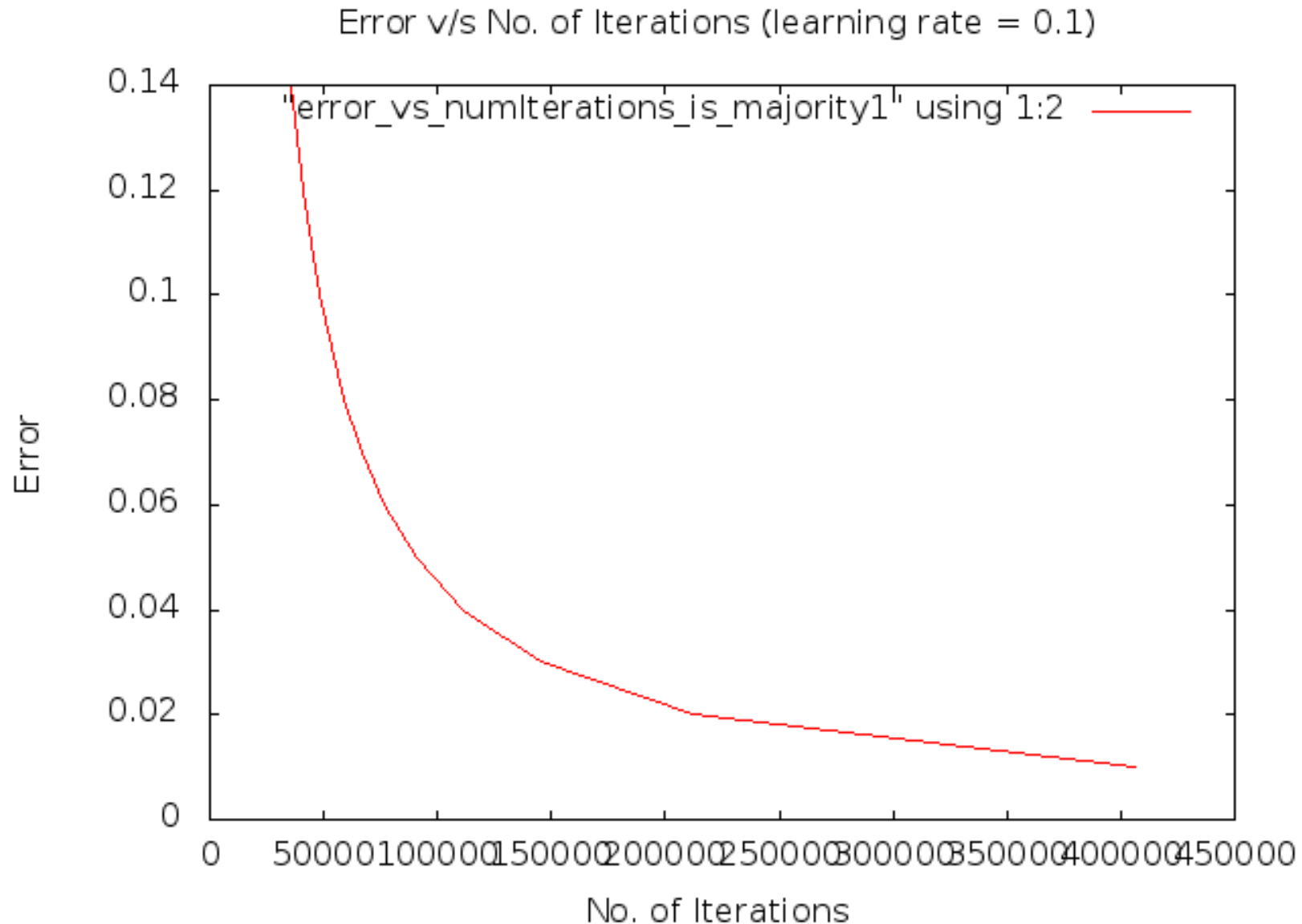


# Effect of Learning Rate (NAND)

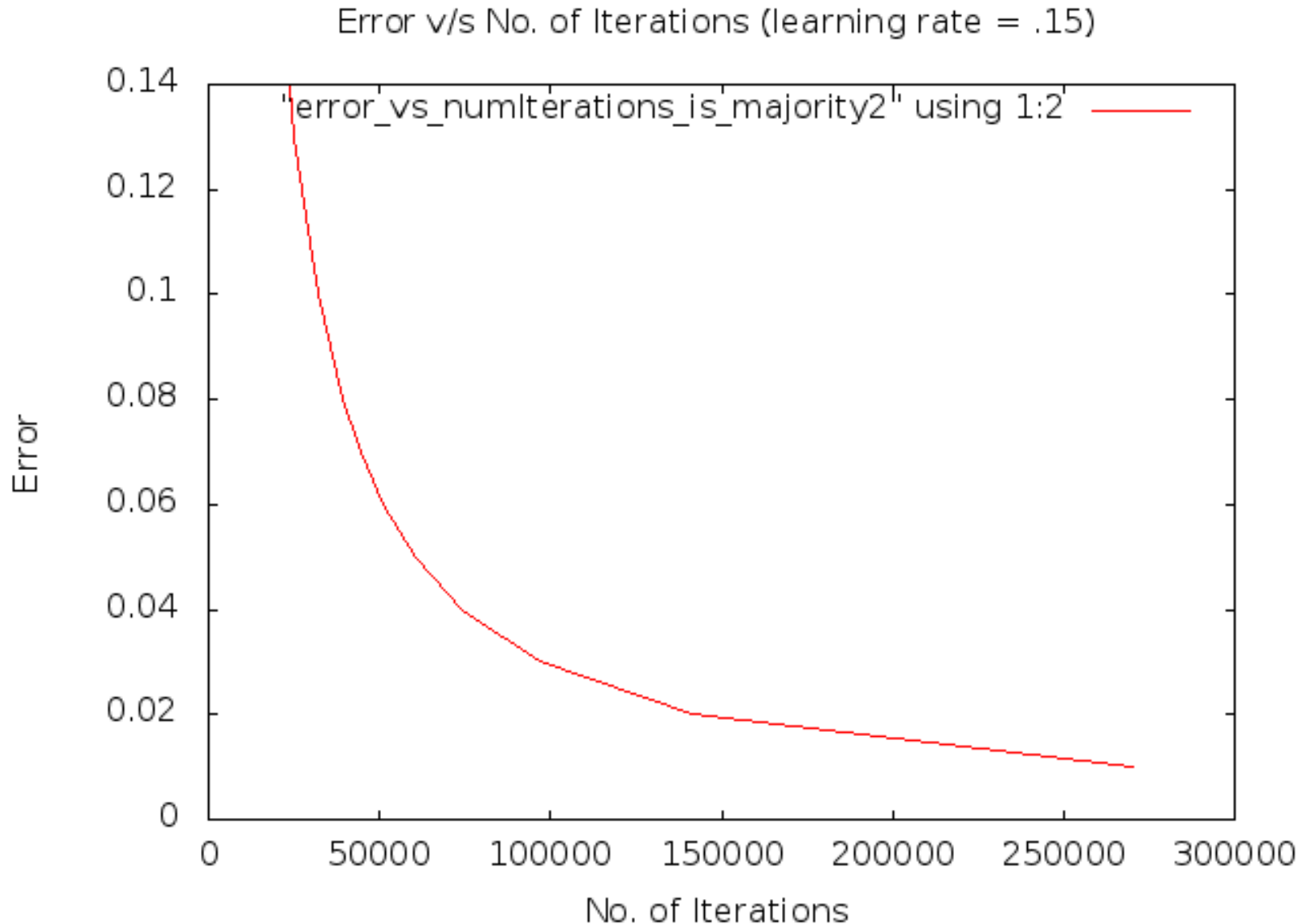




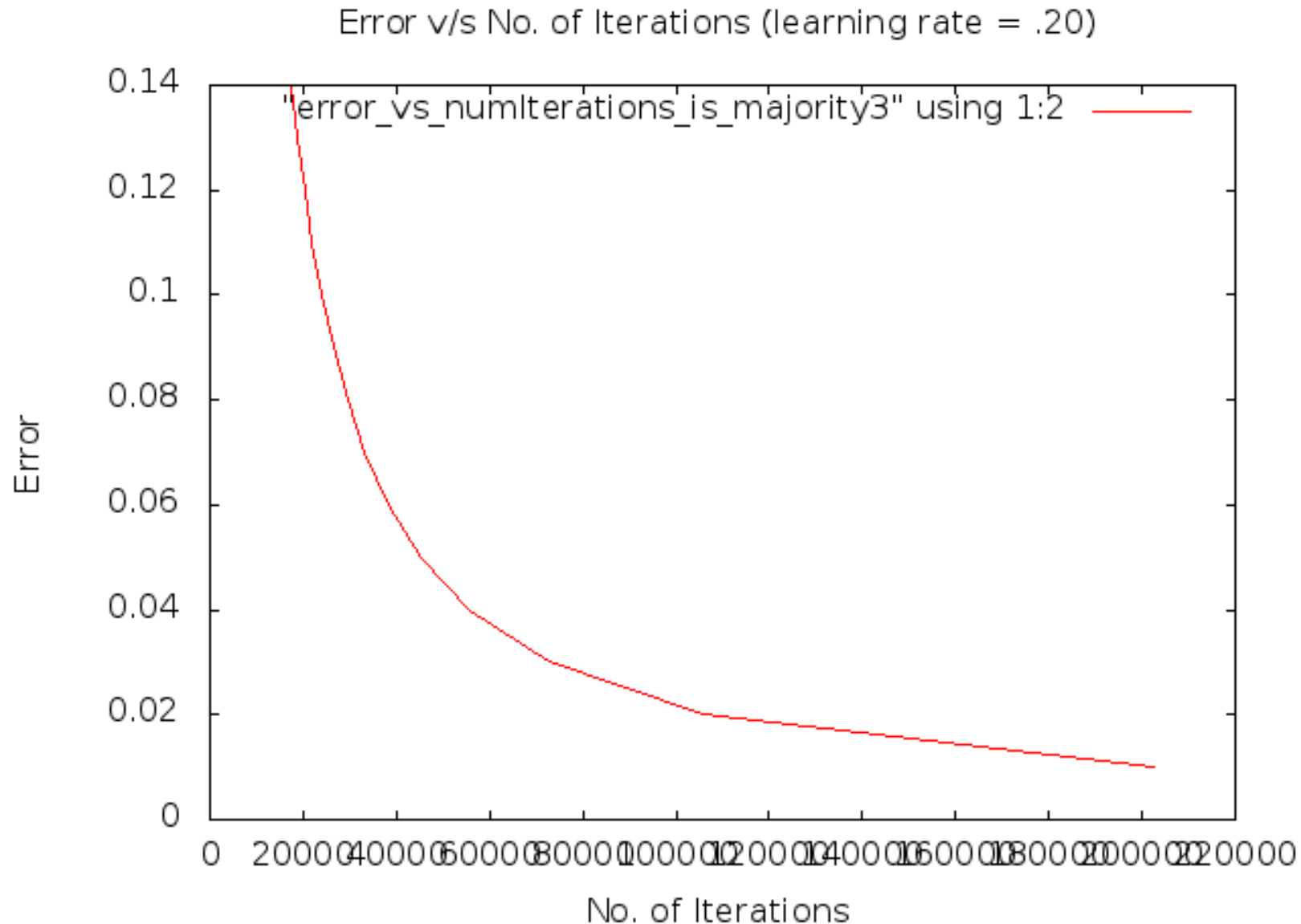
# Effect of Learning Rate (Majority)



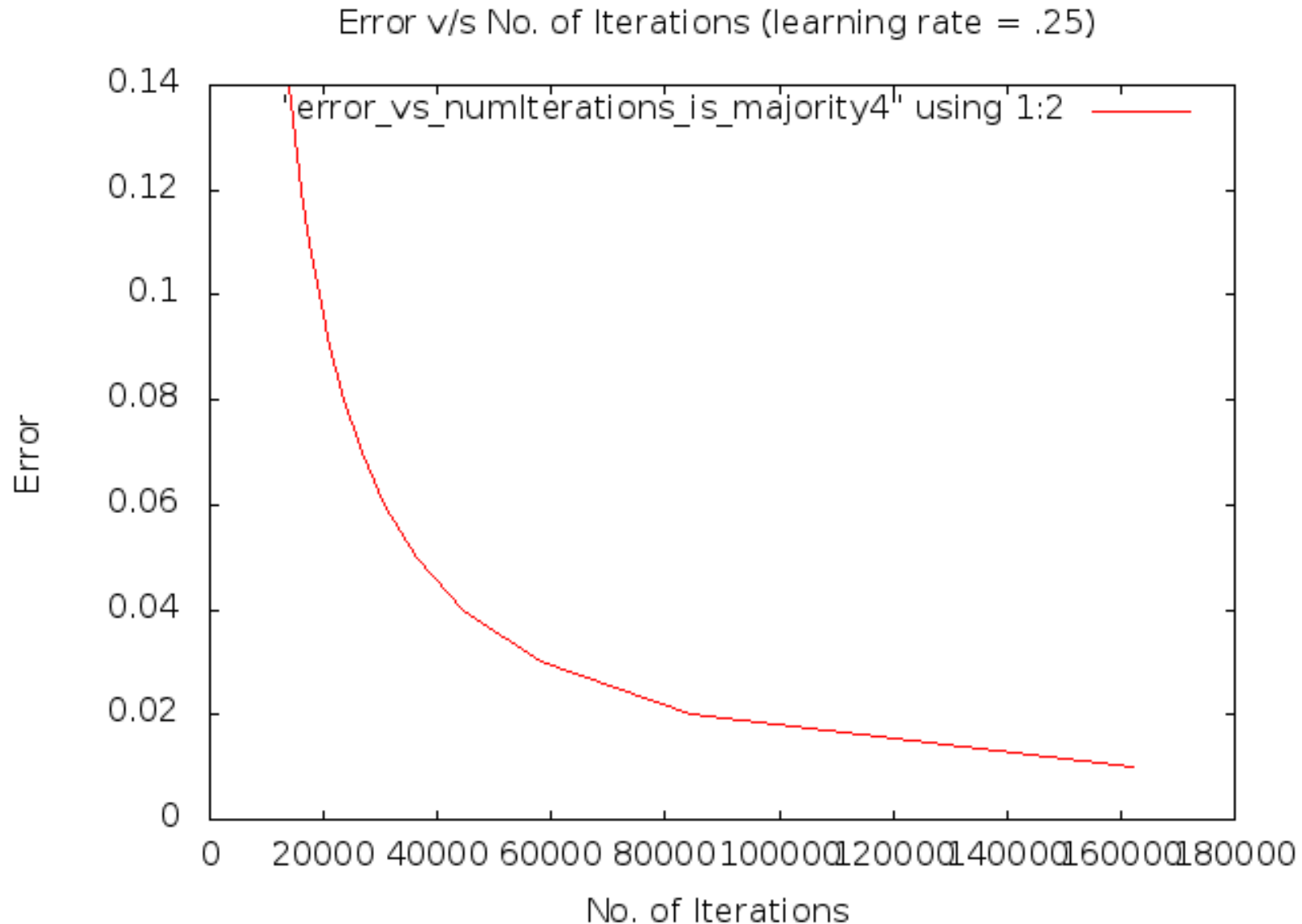
# Effect of Learning Rate (Majority)



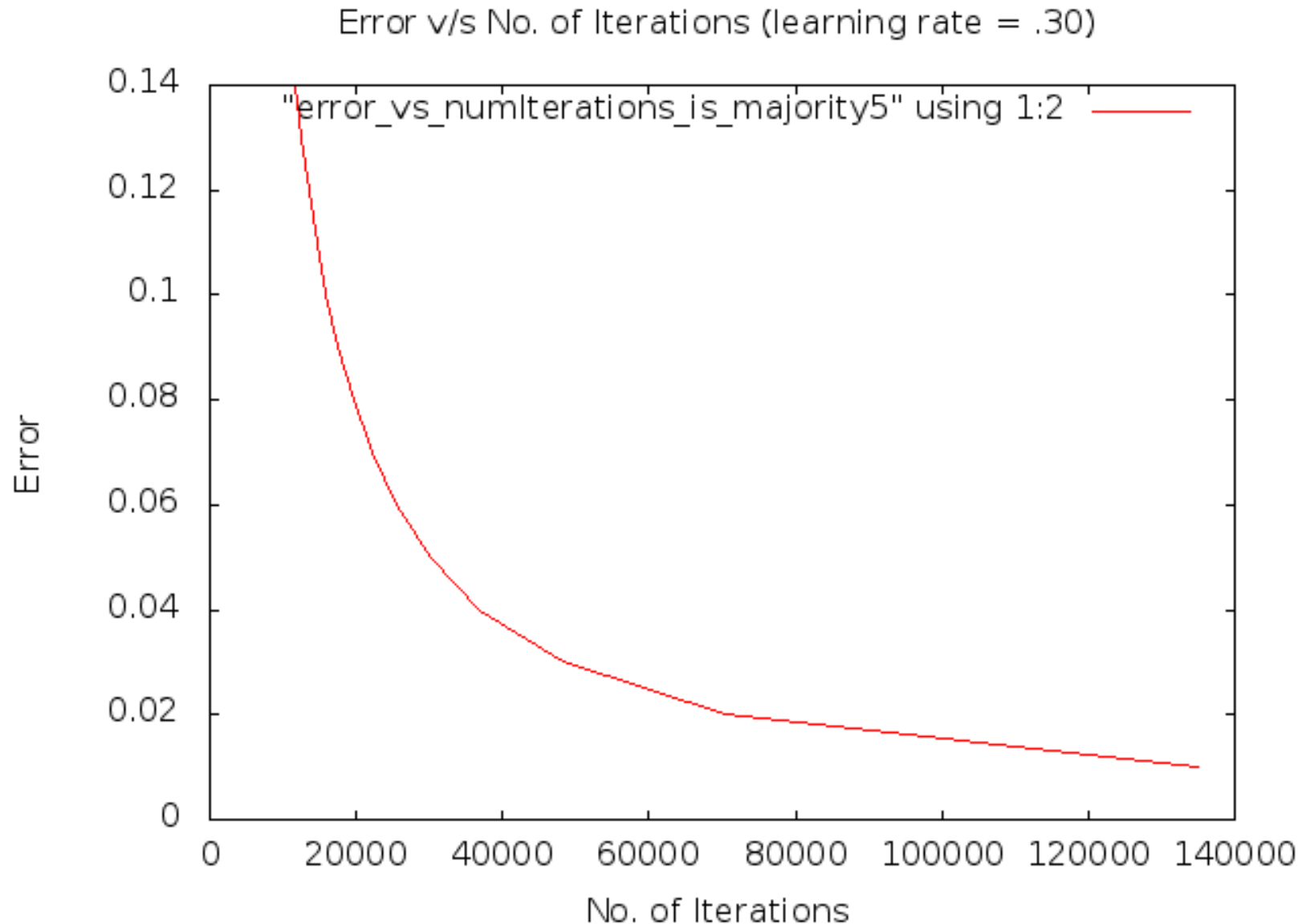
# Effect of Learning Rate (Majority)



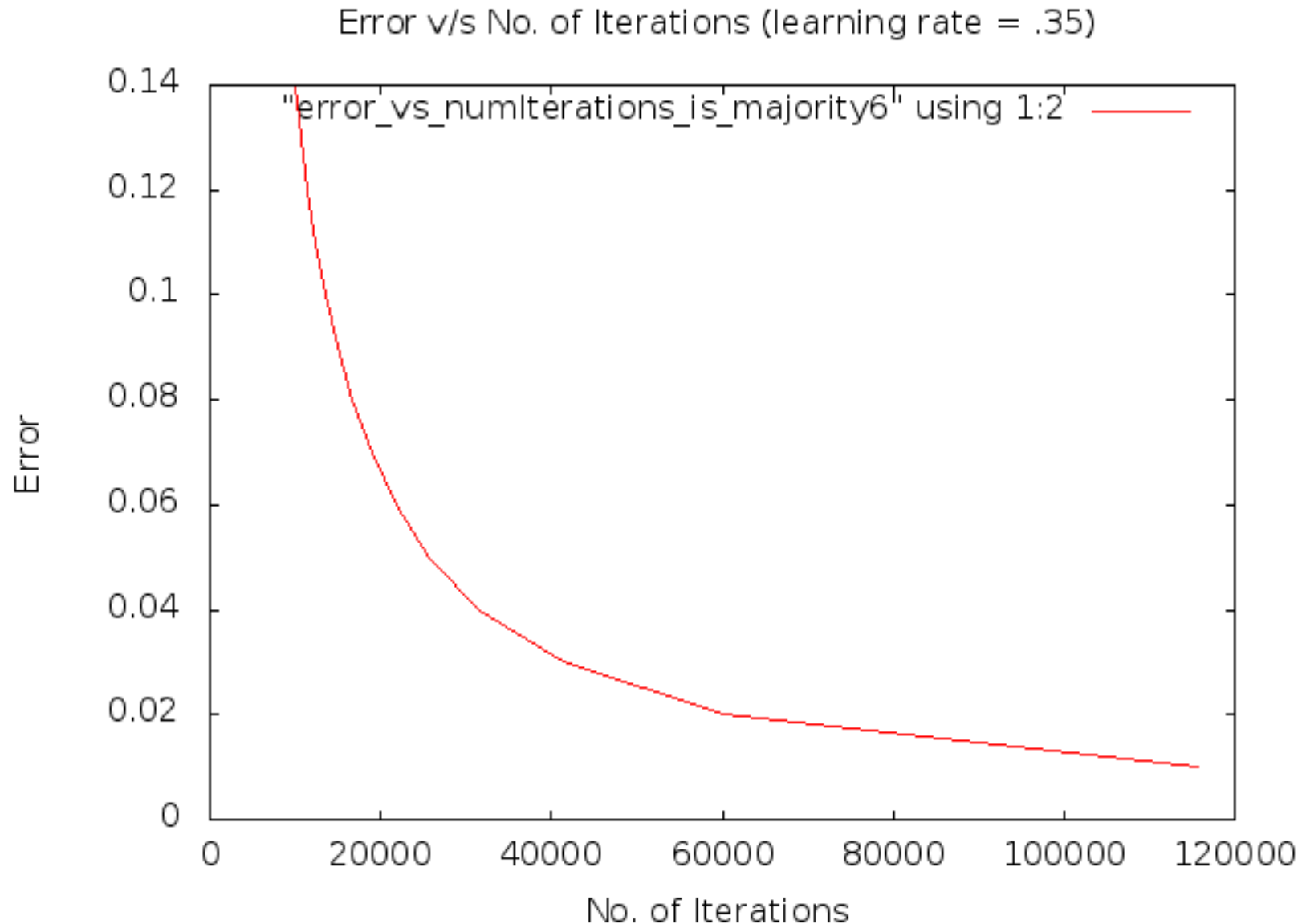
# Effect of Learning Rate (Majority)



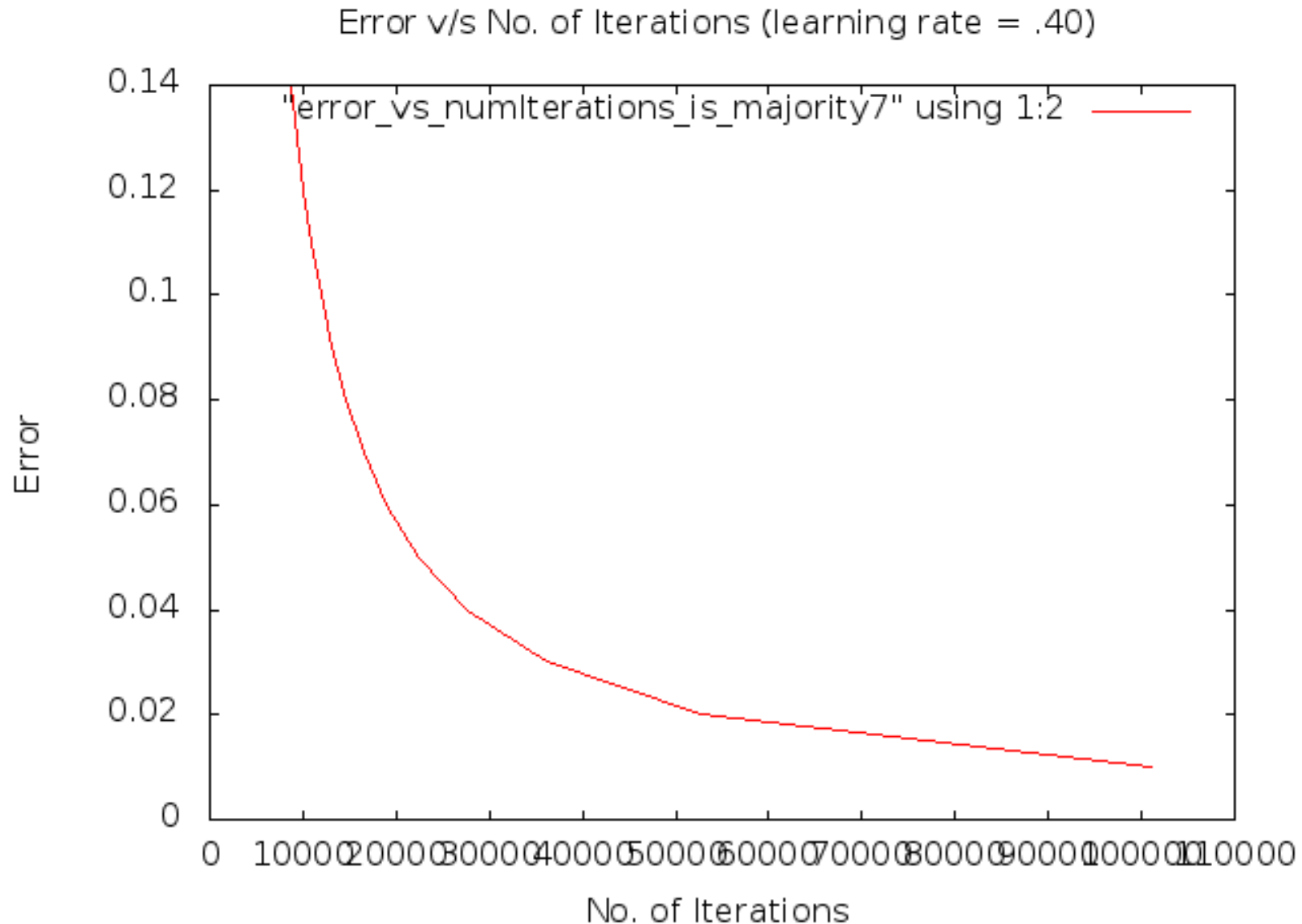
# Effect of Learning Rate (Majority)



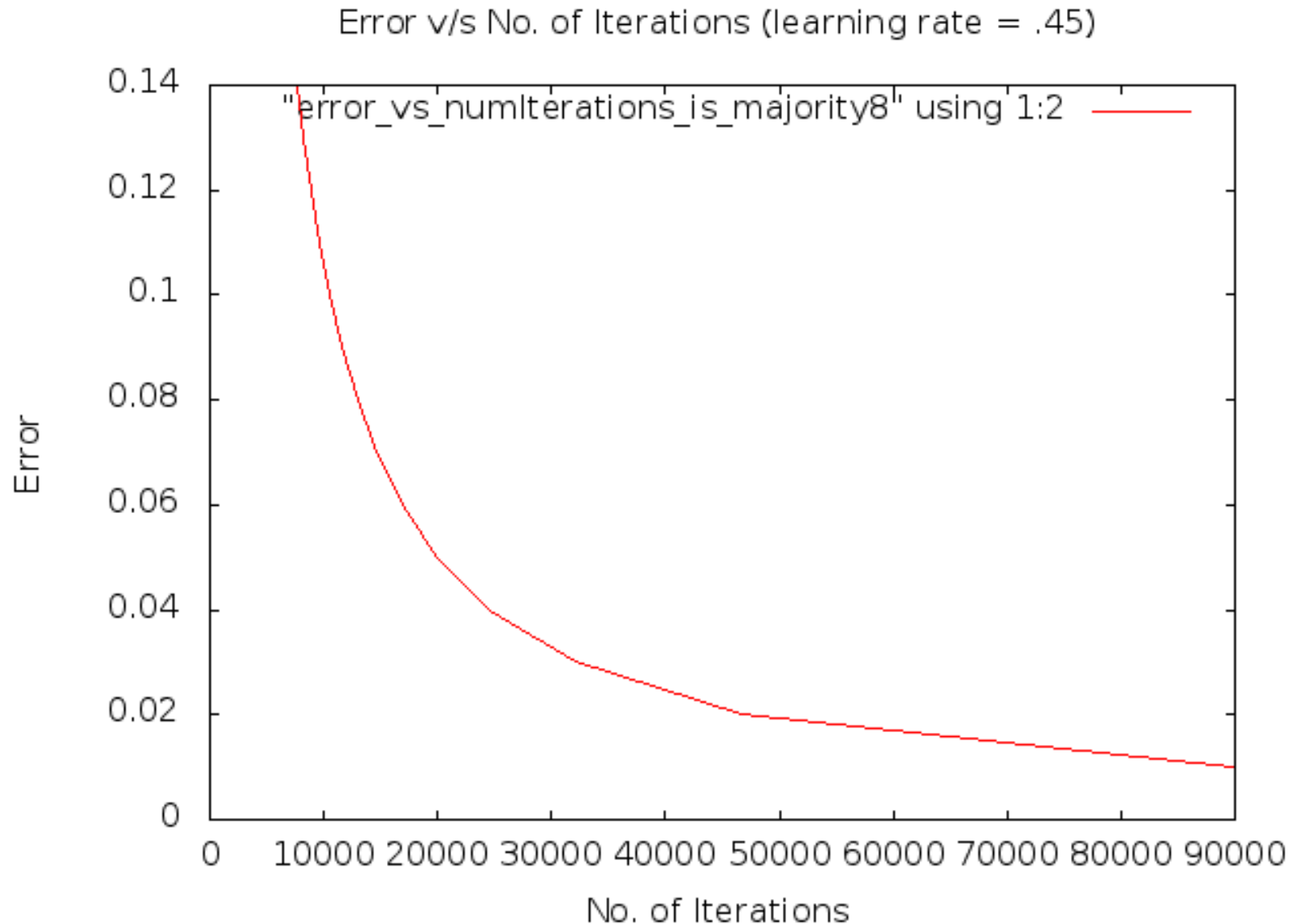
# Effect of Learning Rate (Majority)



# Effect of Learning Rate (Majority)

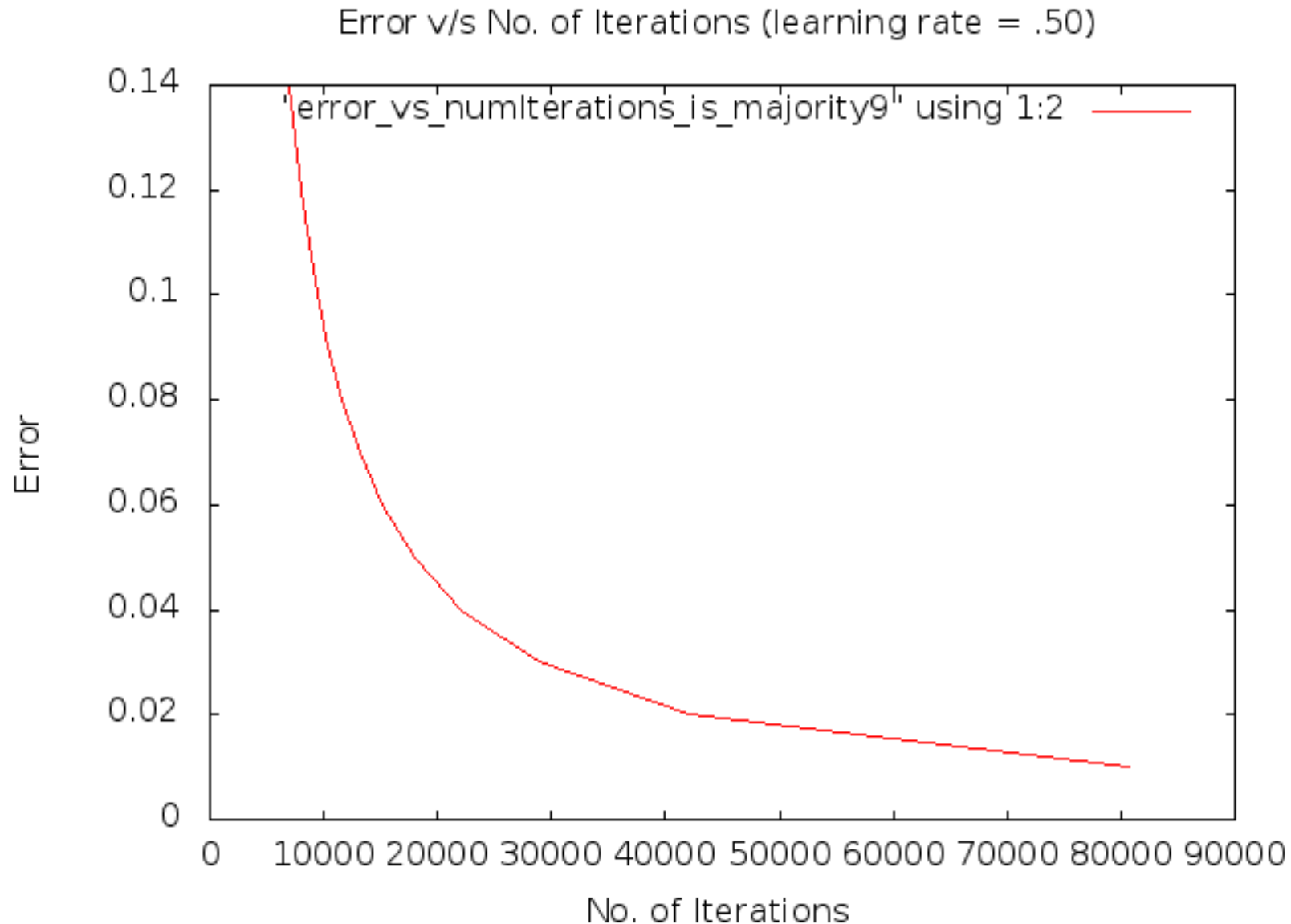


# Effect of Learning Rate (Majority)

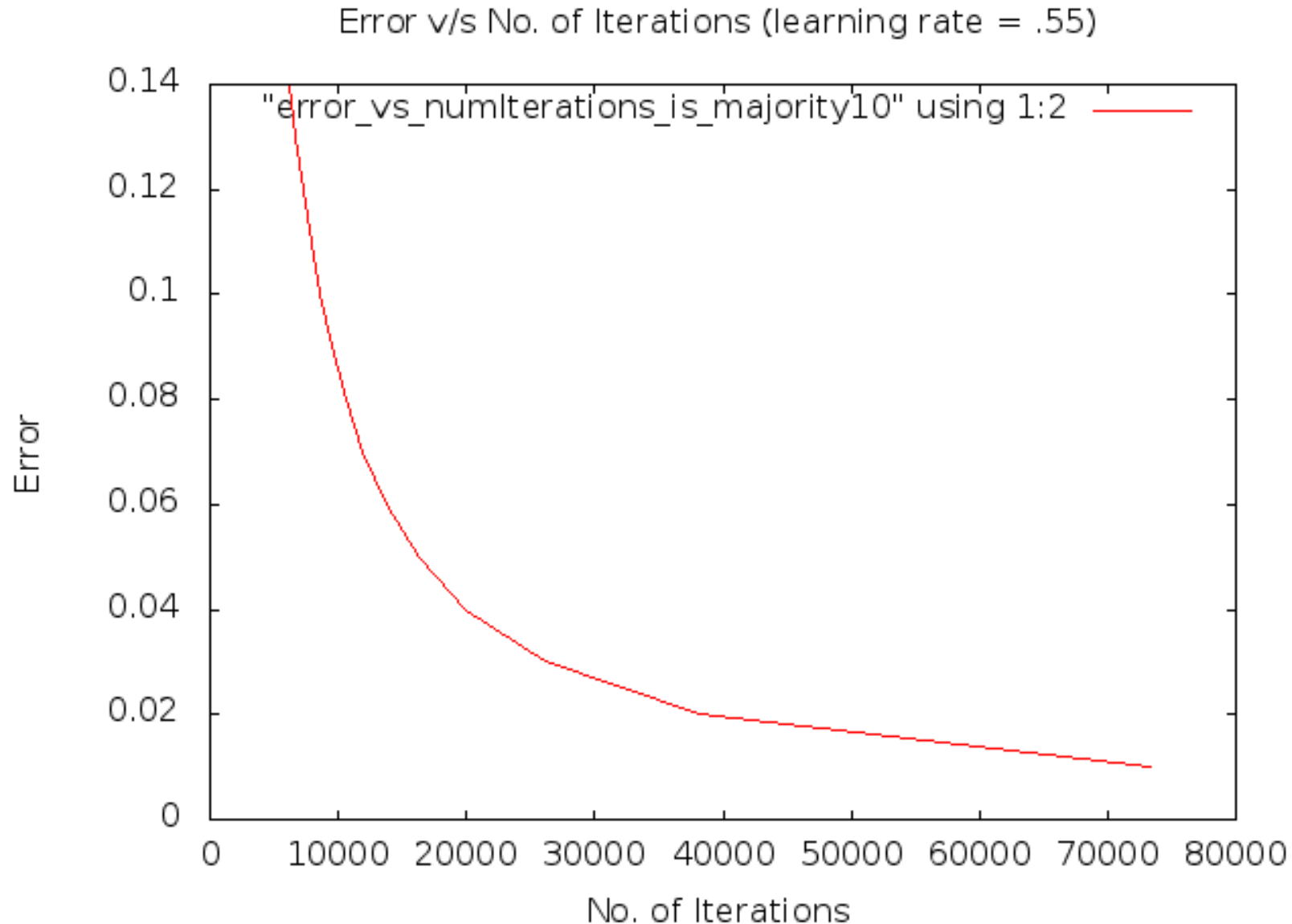




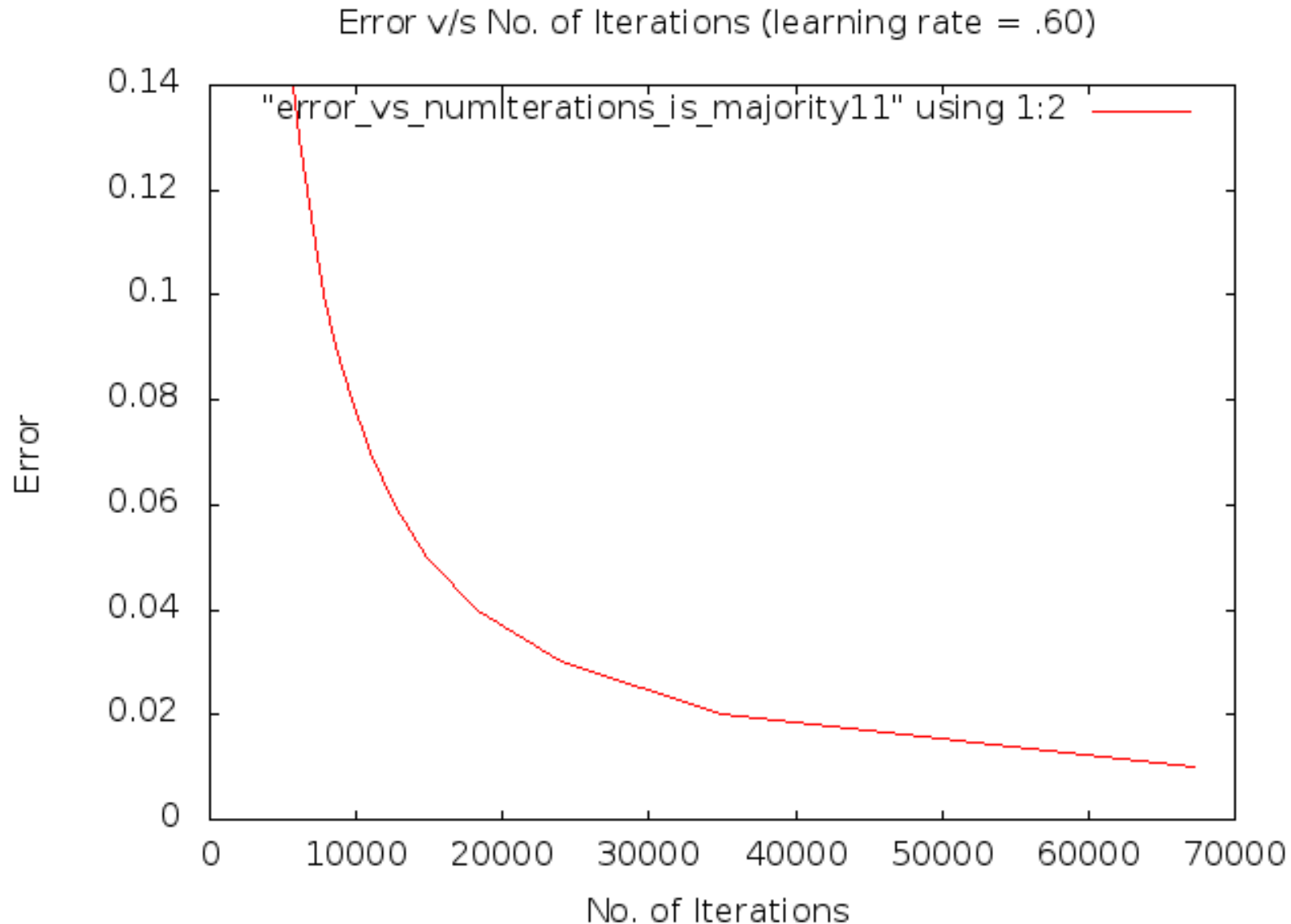
# Effect of Learning Rate (Majority)



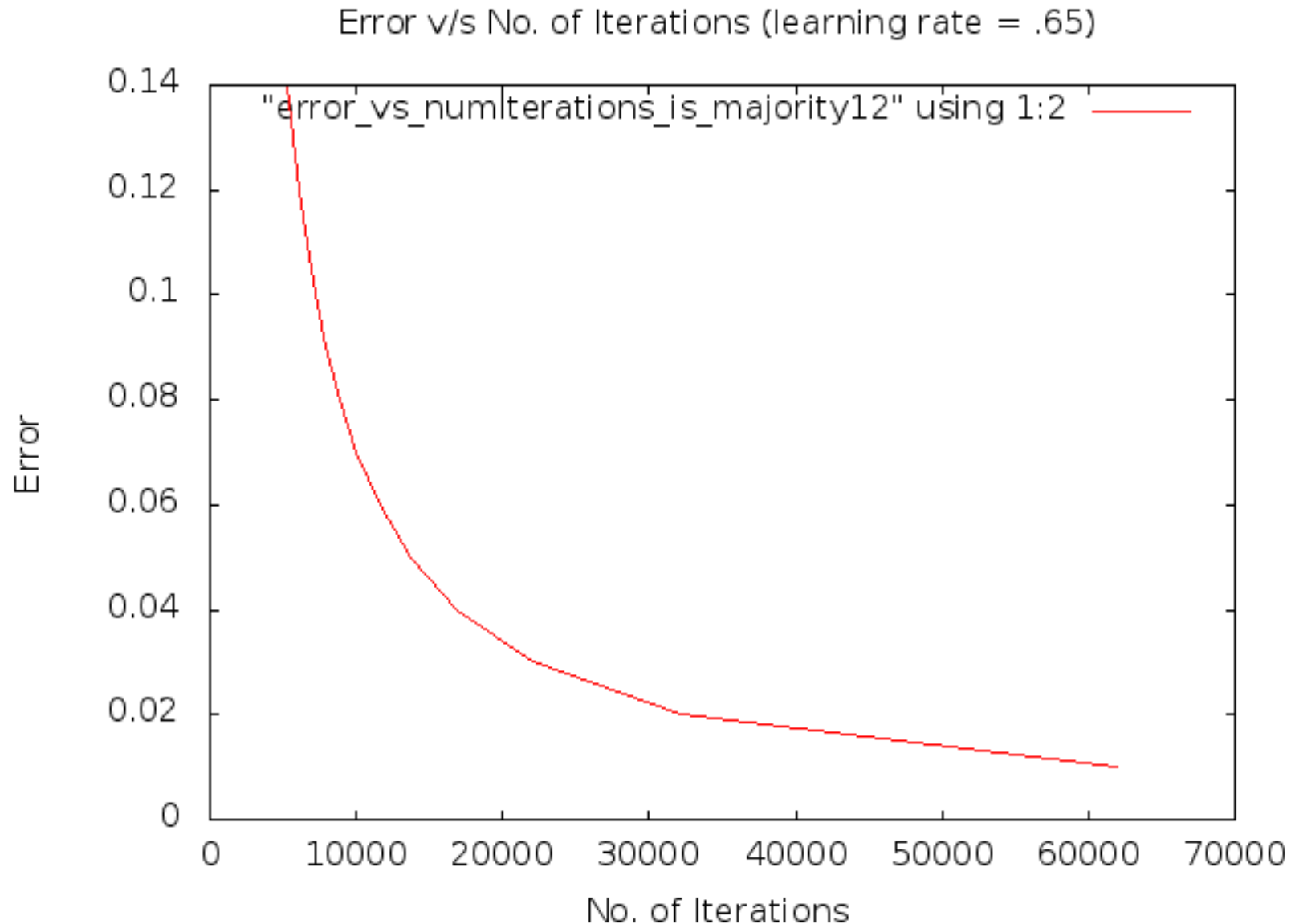
# Effect of Learning Rate (Majority)



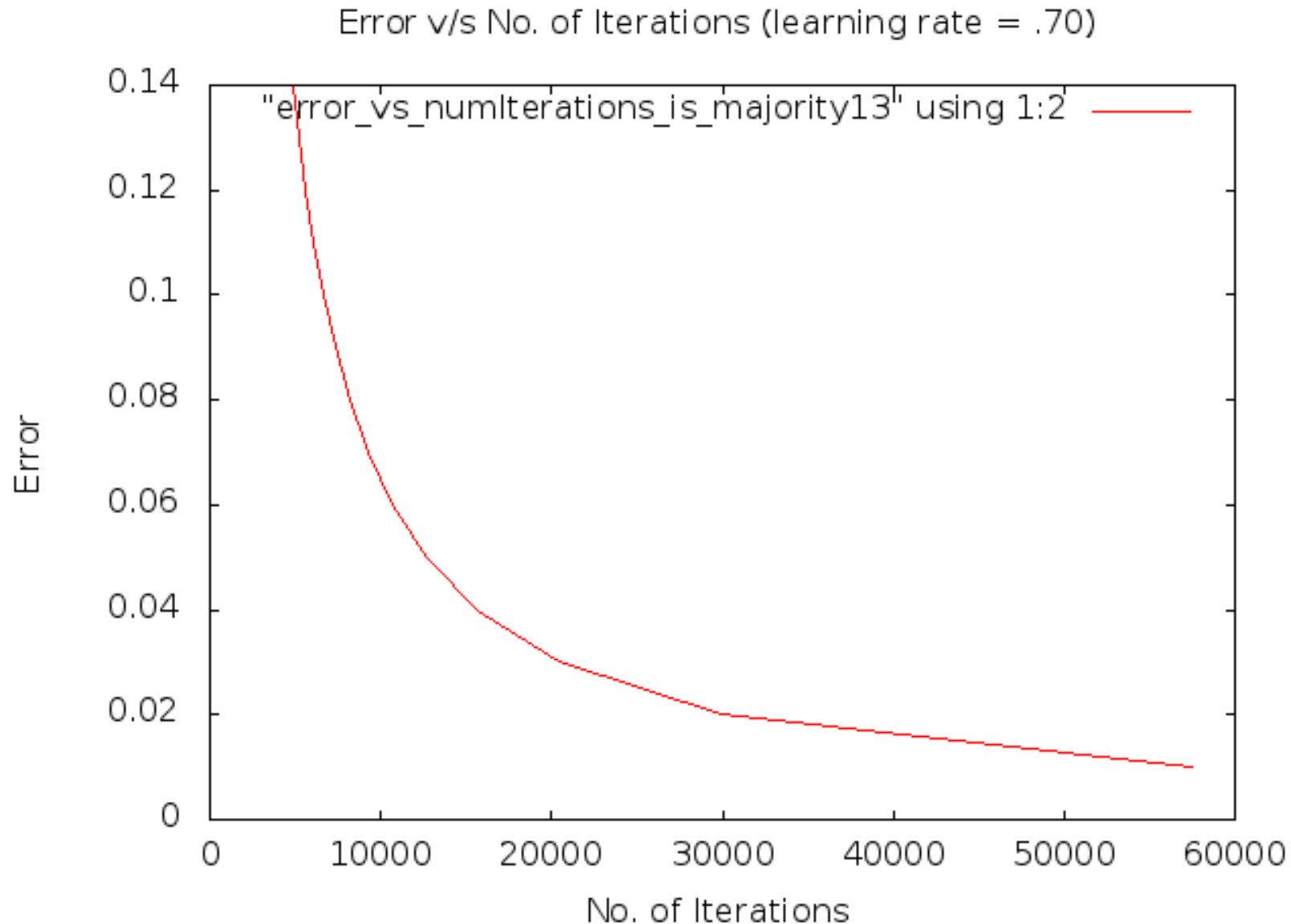
# Effect of Learning Rate (Majority)



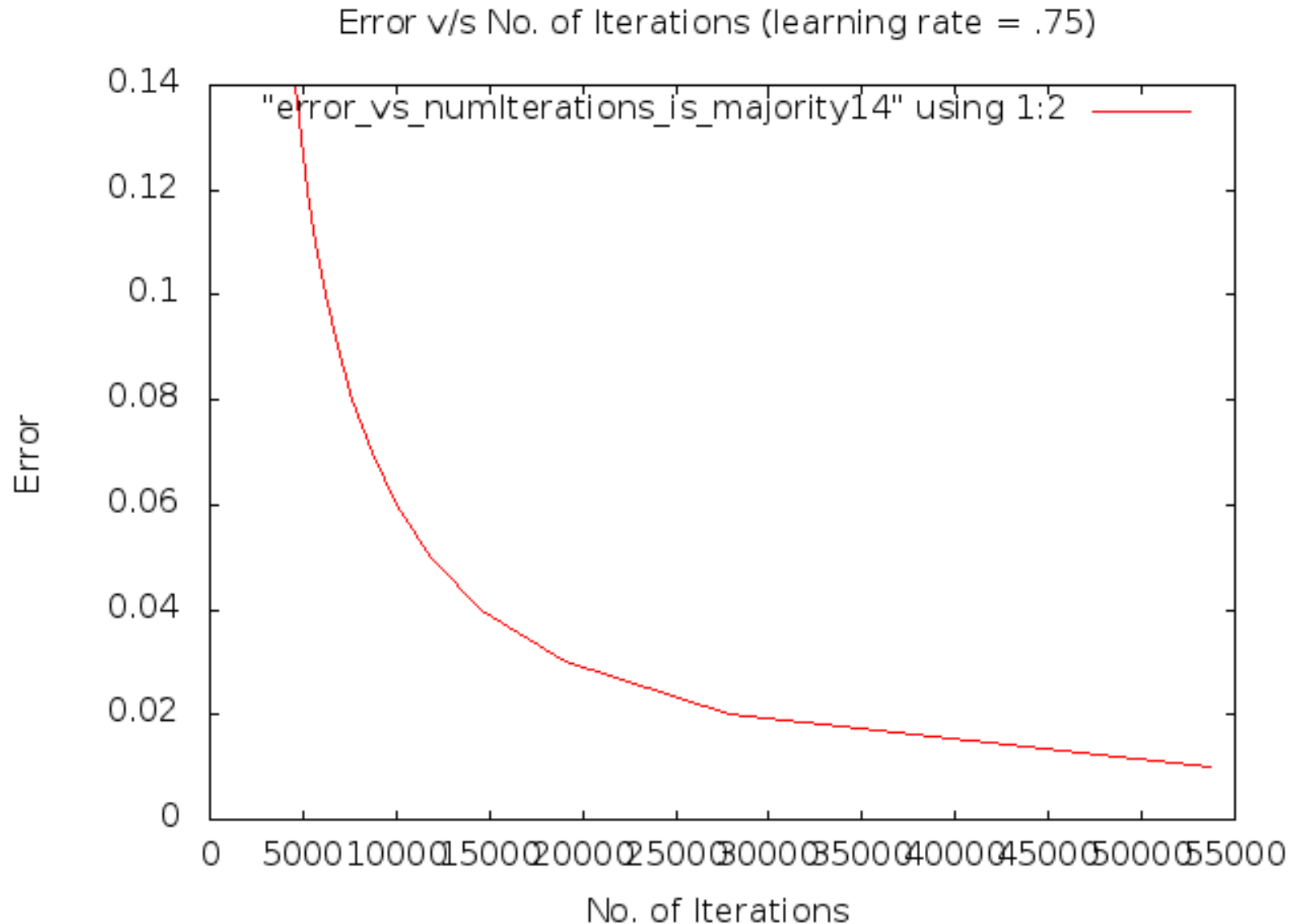
# Effect of Learning Rate (Majority)



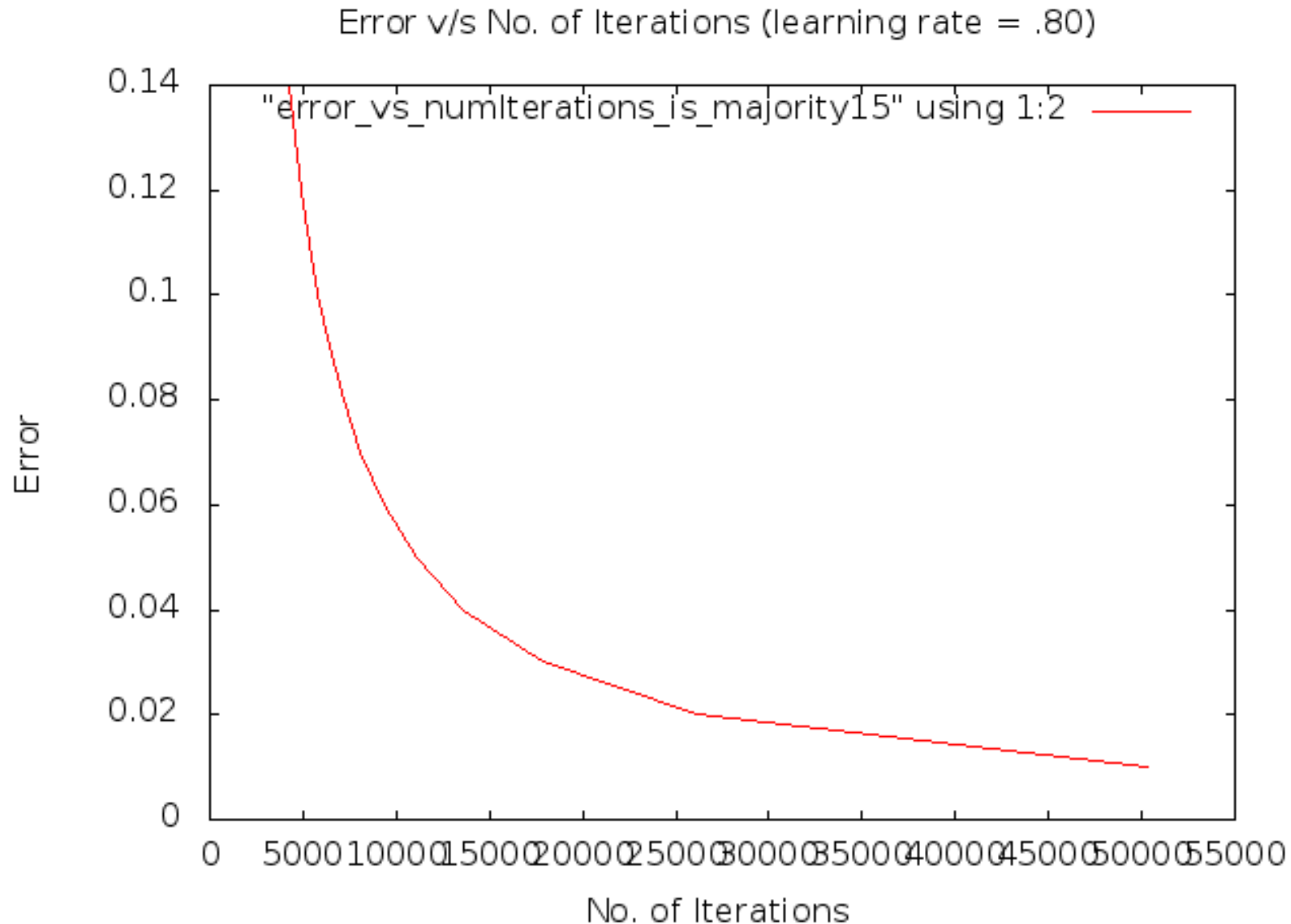
# Effect of Learning Rate (Majority)



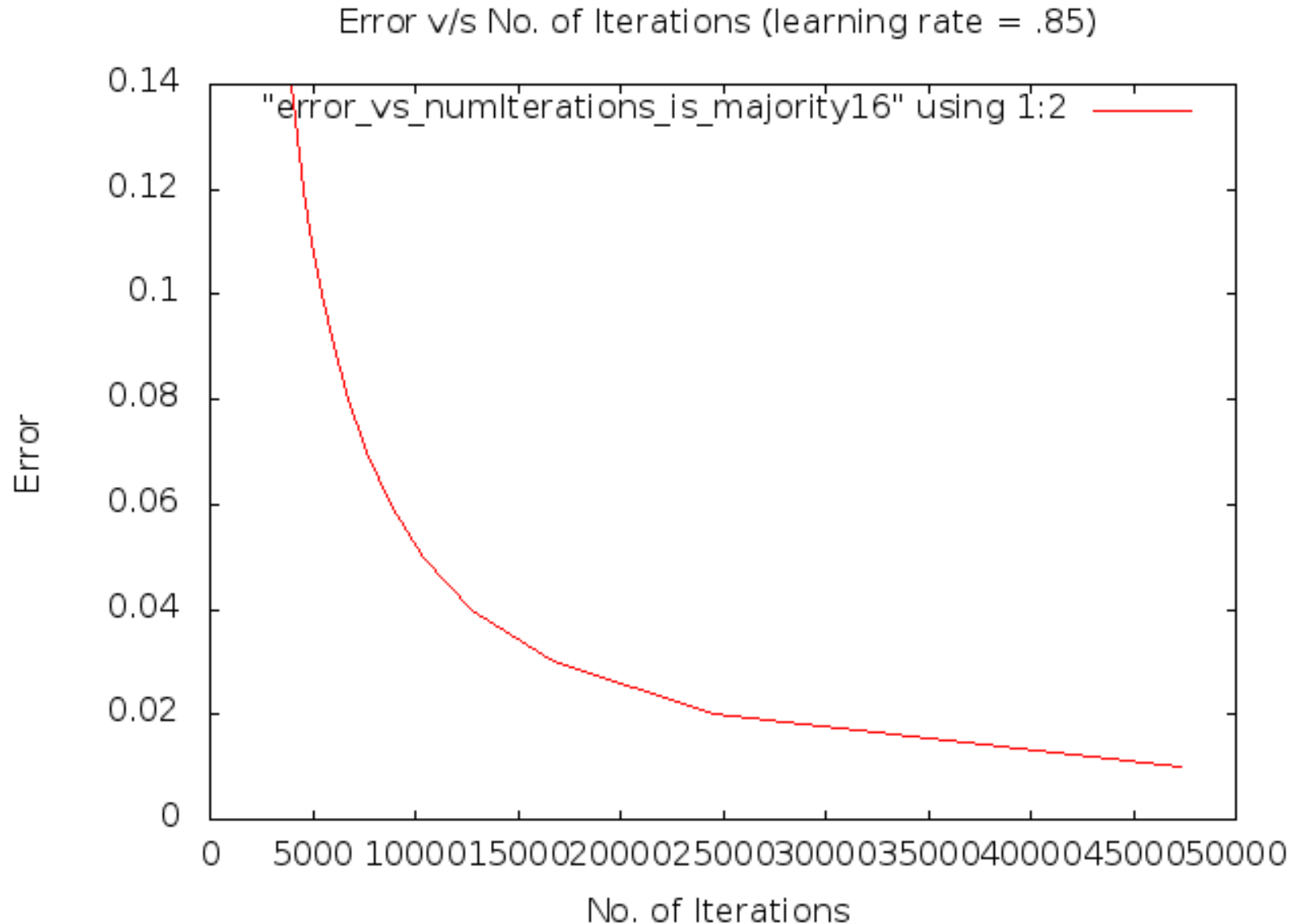
# Effect of Learning Rate (Majority)



# Effect of Learning Rate (Majority)

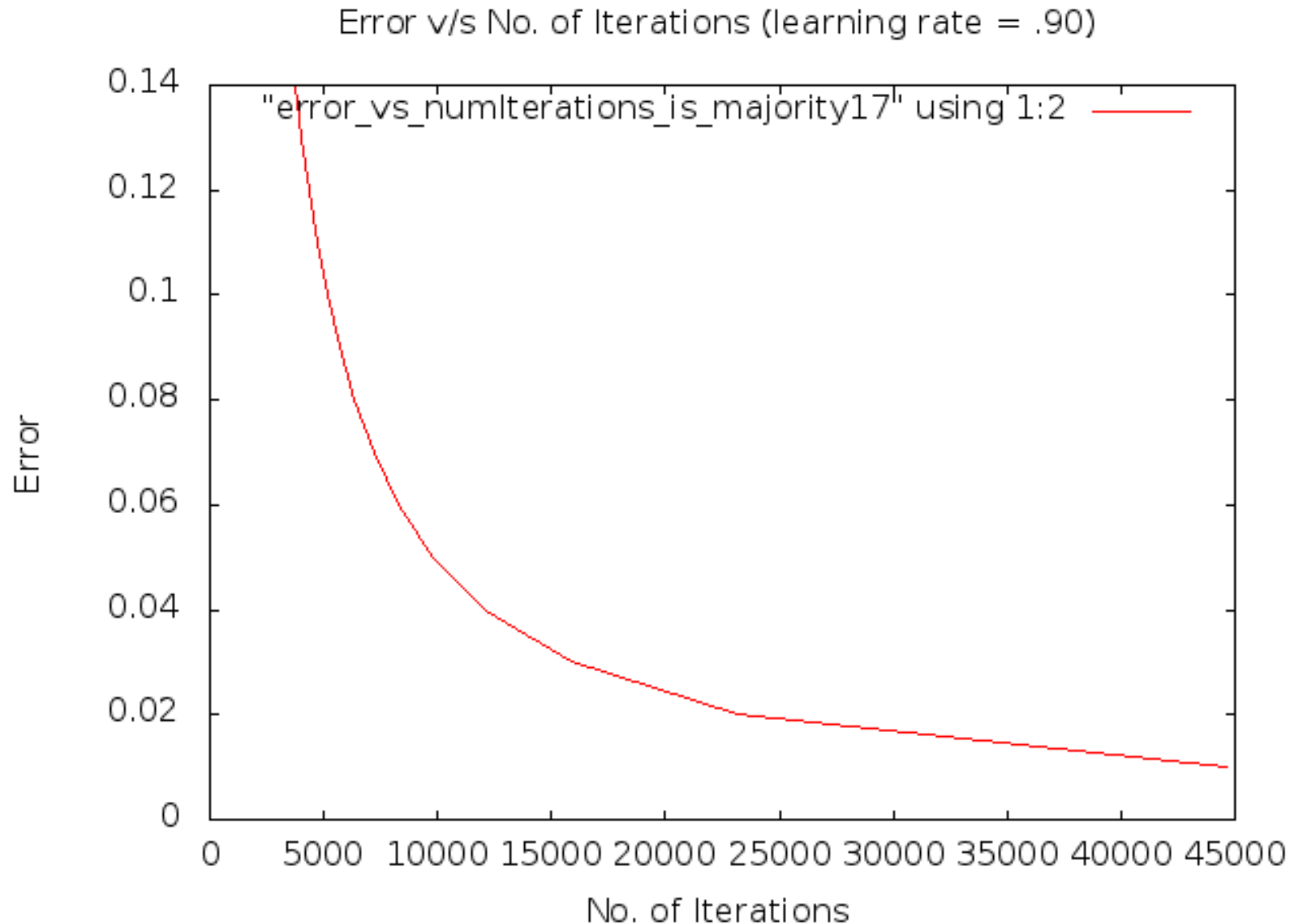


# Effect of Learning Rate (Majority)

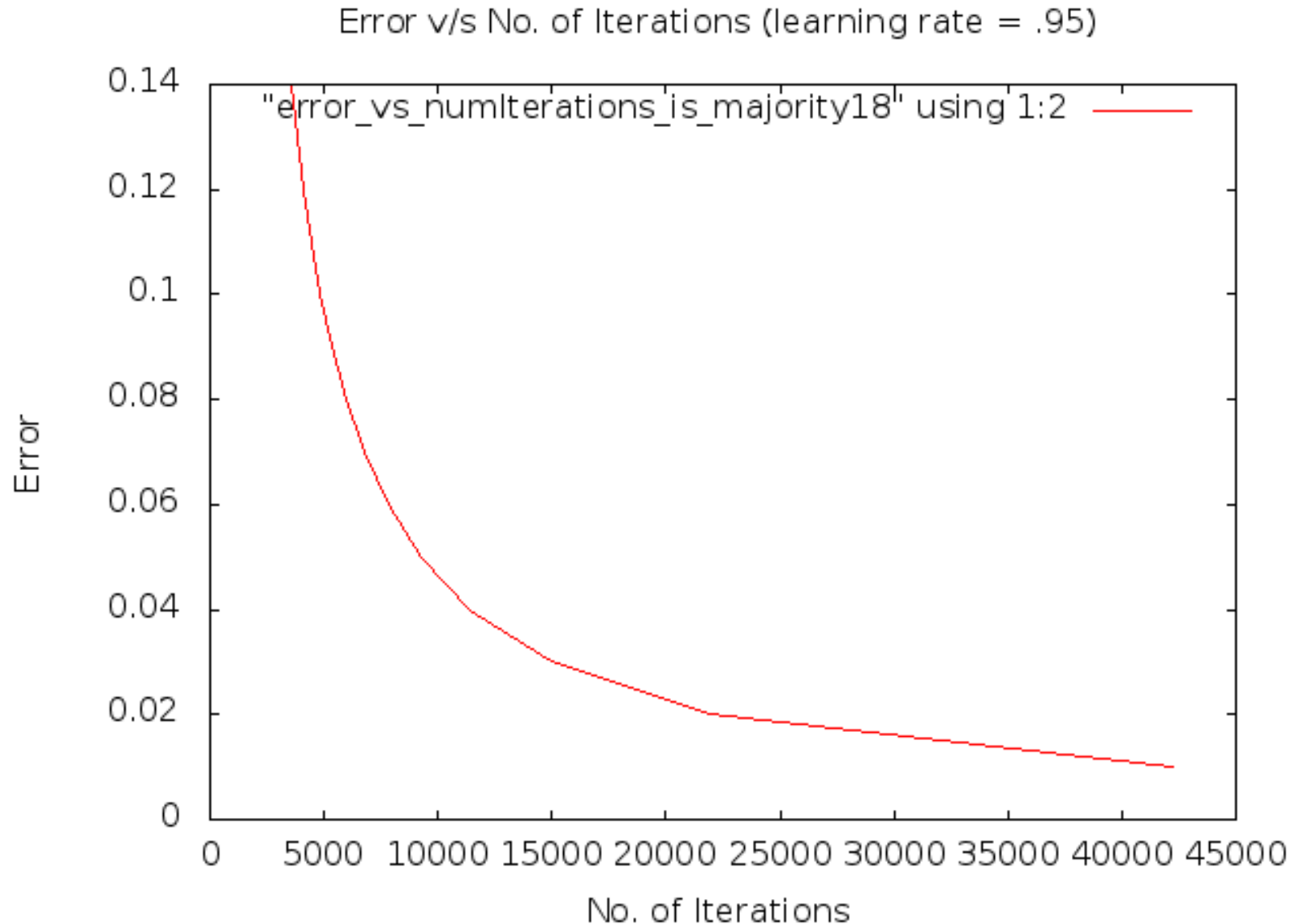




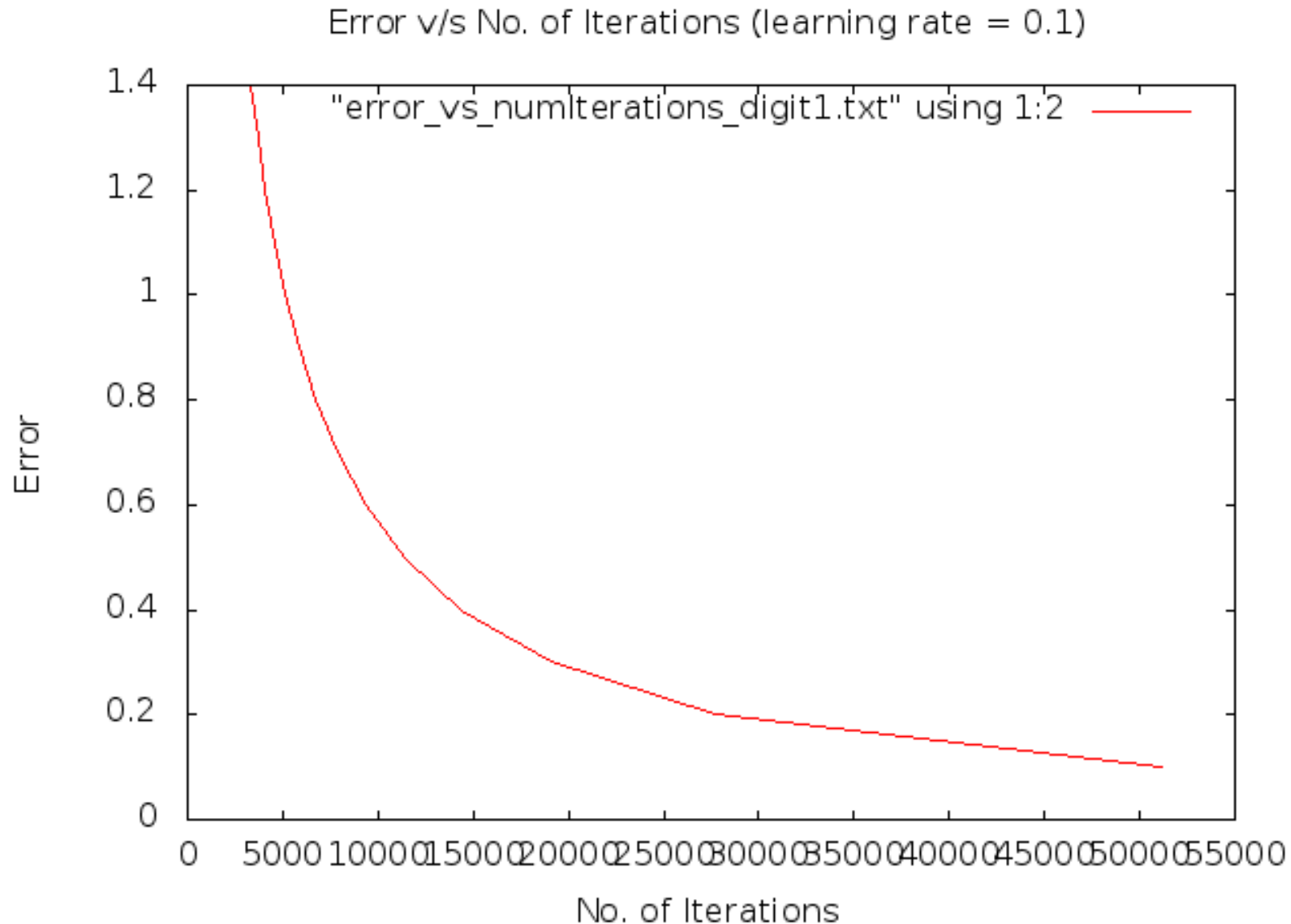
# Effect of Learning Rate (Majority)



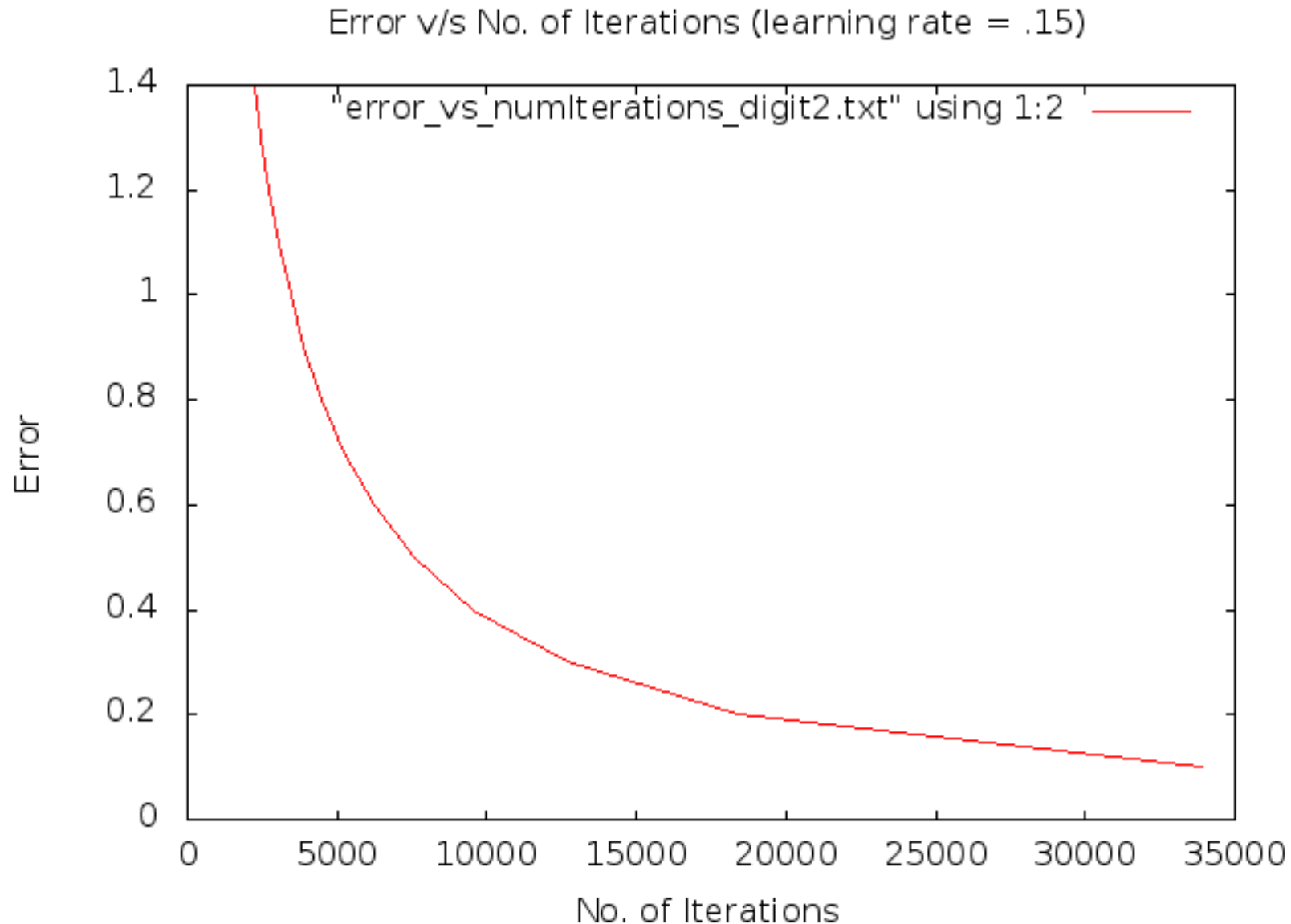
# Effect of Learning Rate (Majority)



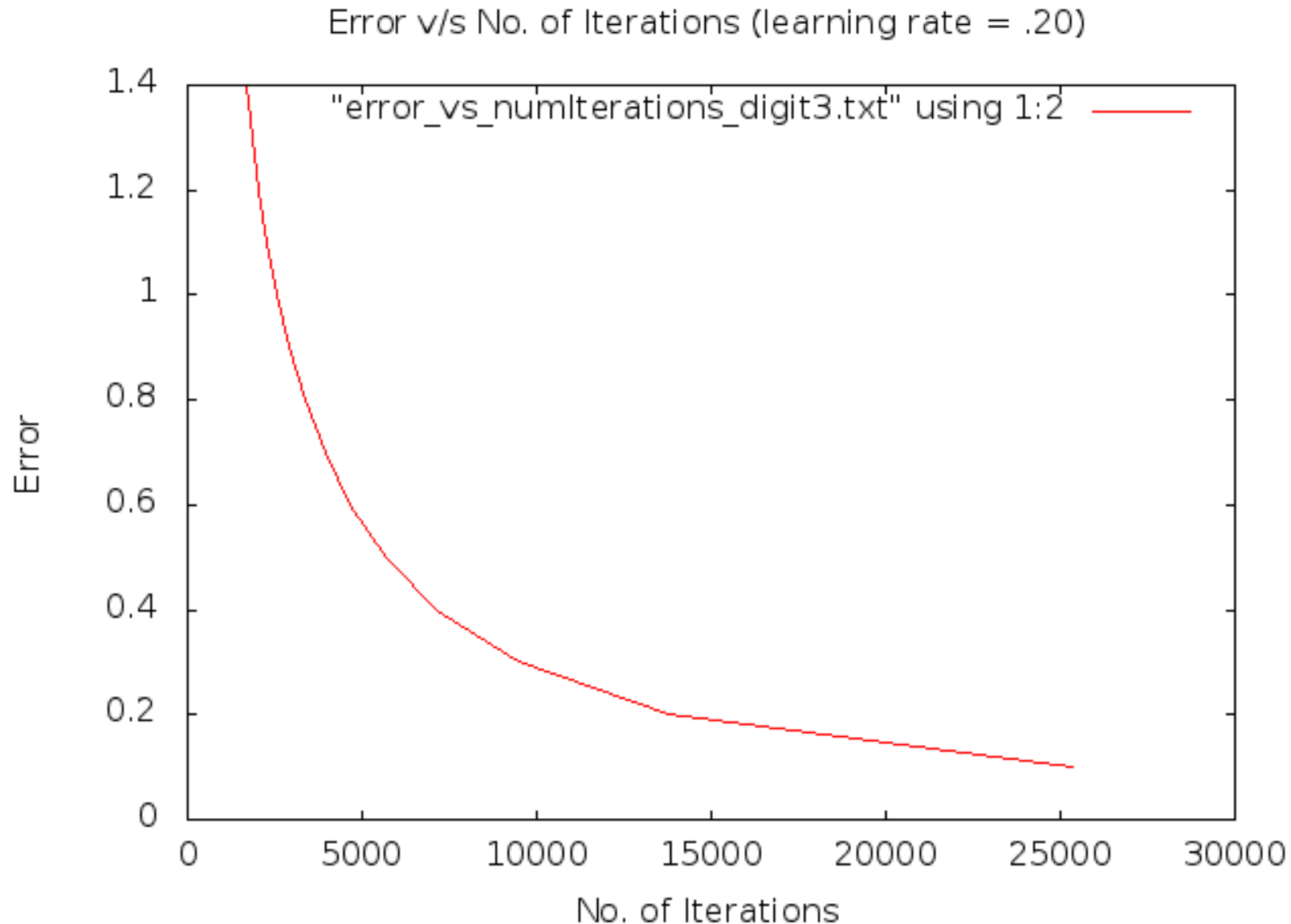
# Effect of Learning Rate (Digit Recogniser)



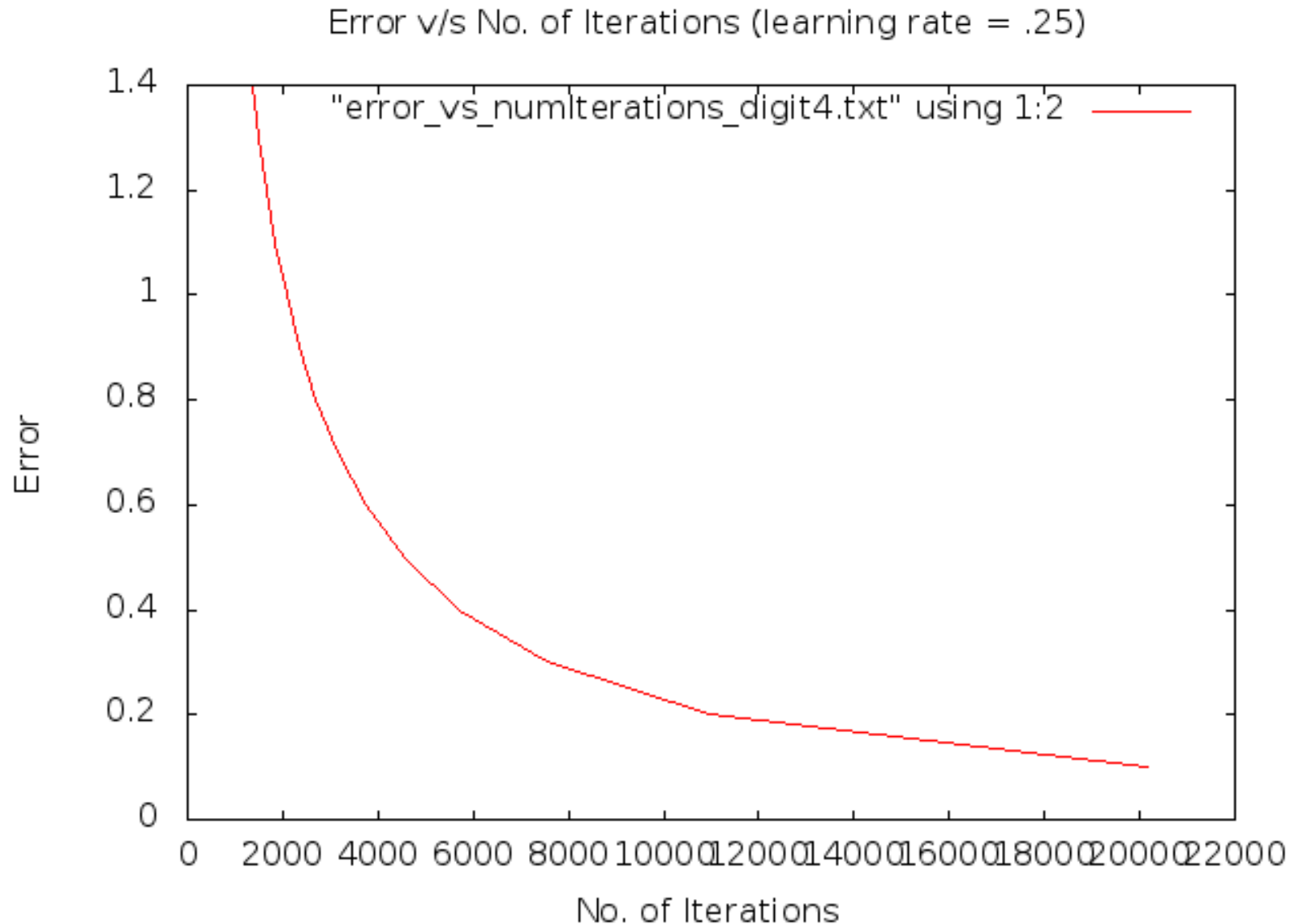
# Effect of Learning Rate (Digit Recogniser)



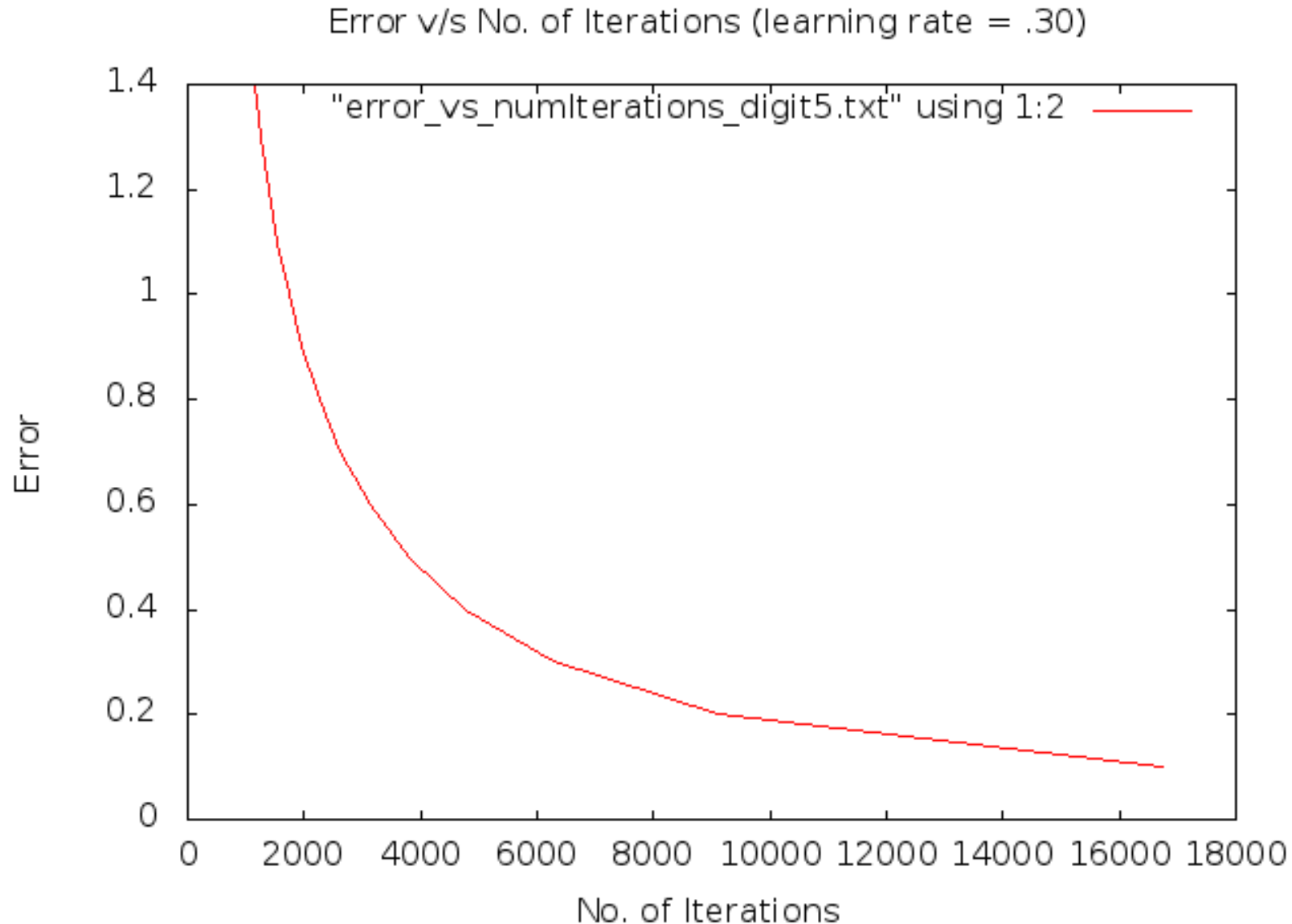
# Effect of Learning Rate (Digit Recogniser)



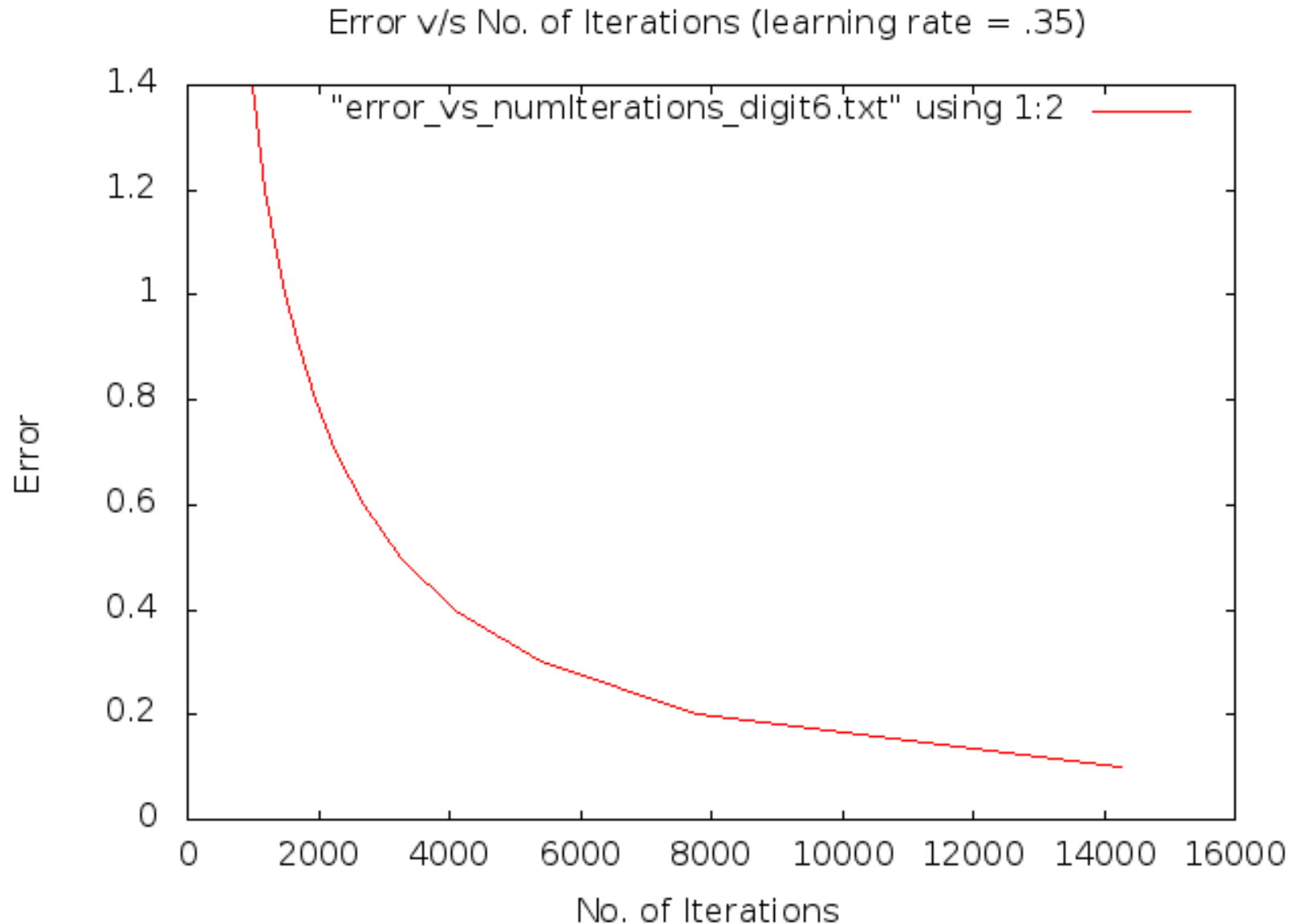
# Effect of Learning Rate (Digit Recogniser)



# Effect of Learning Rate (Digit Recogniser)

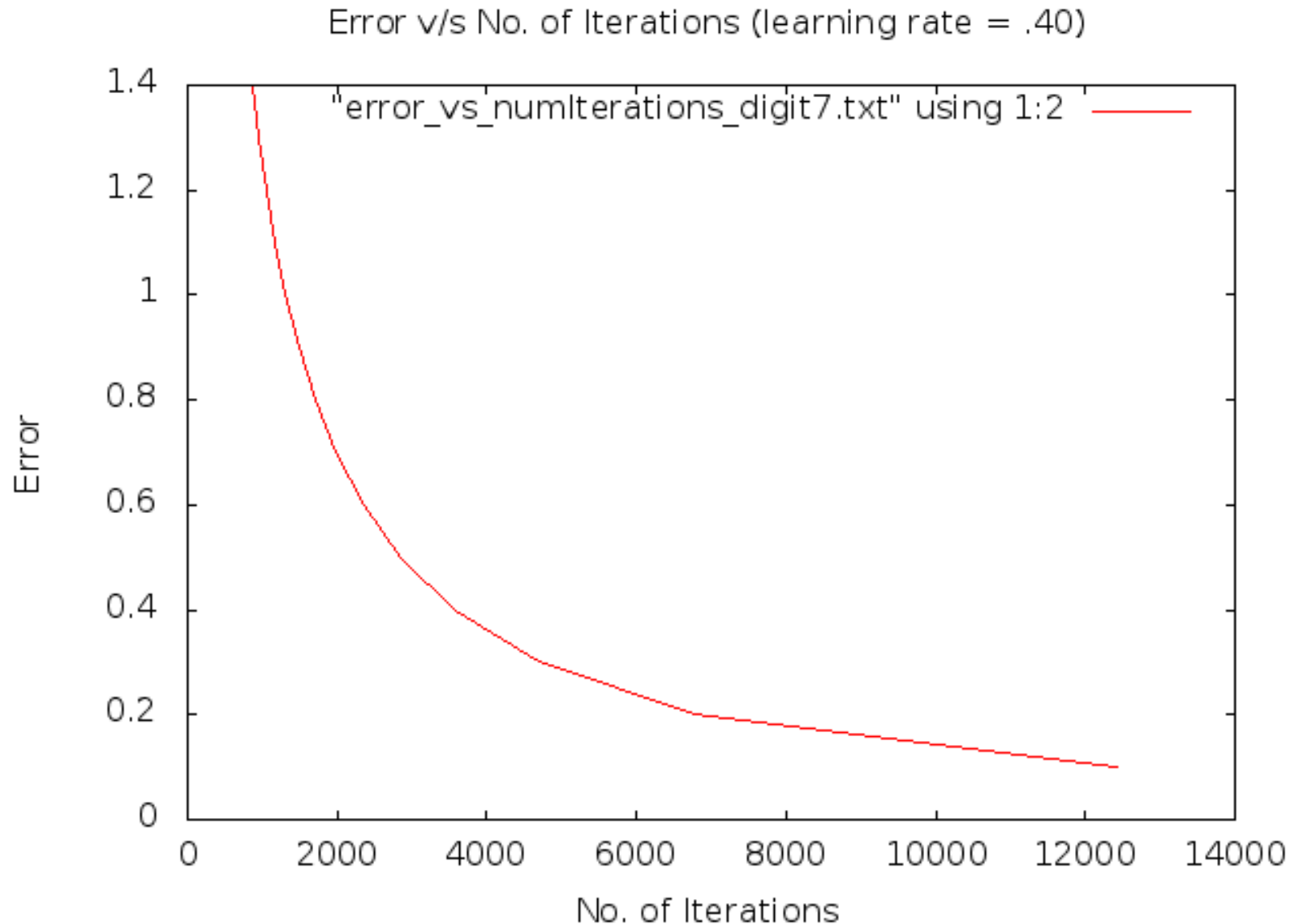


# Effect of Learning Rate (Digit Recogniser)

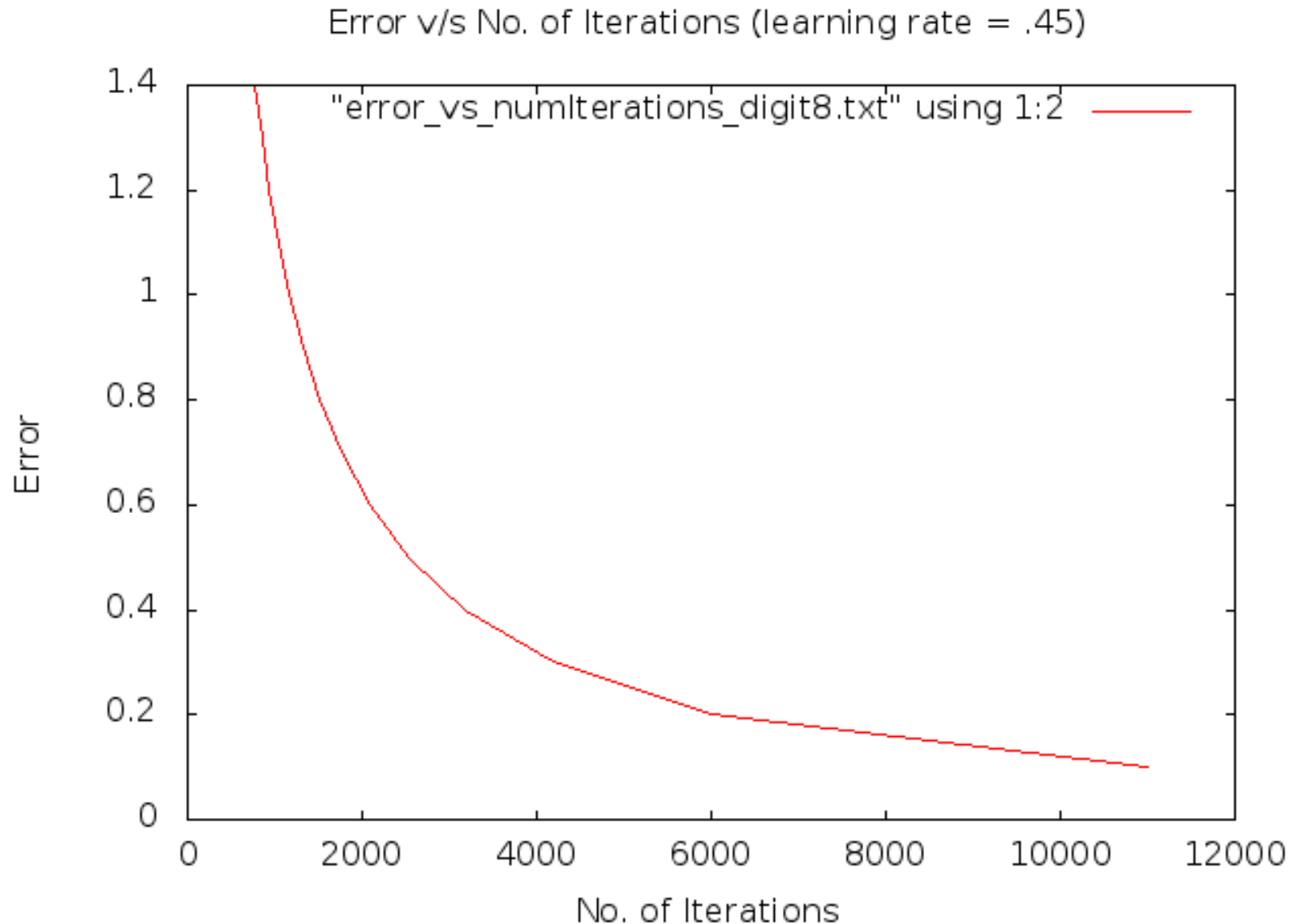




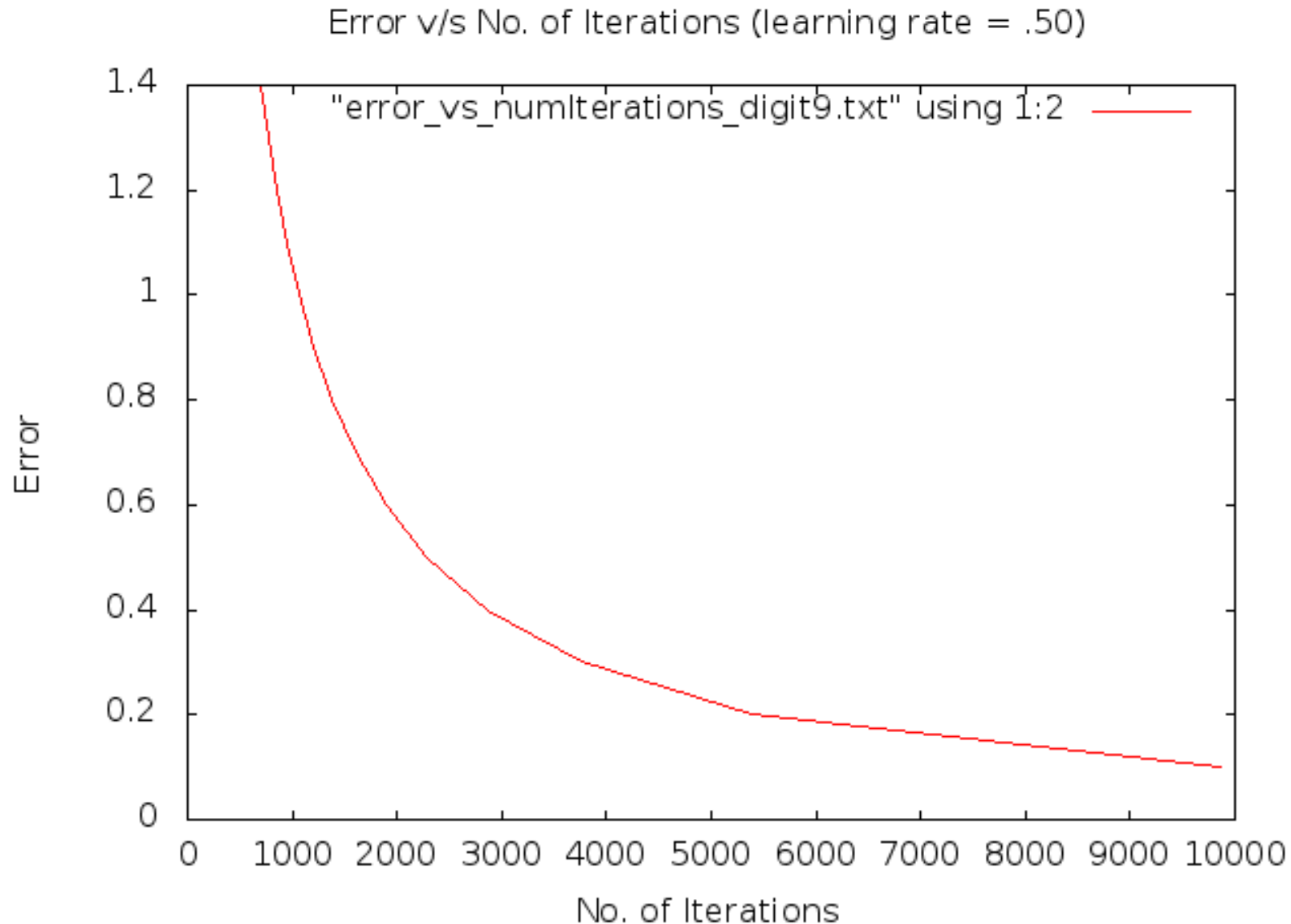
# Effect of Learning Rate (Digit Recogniser)



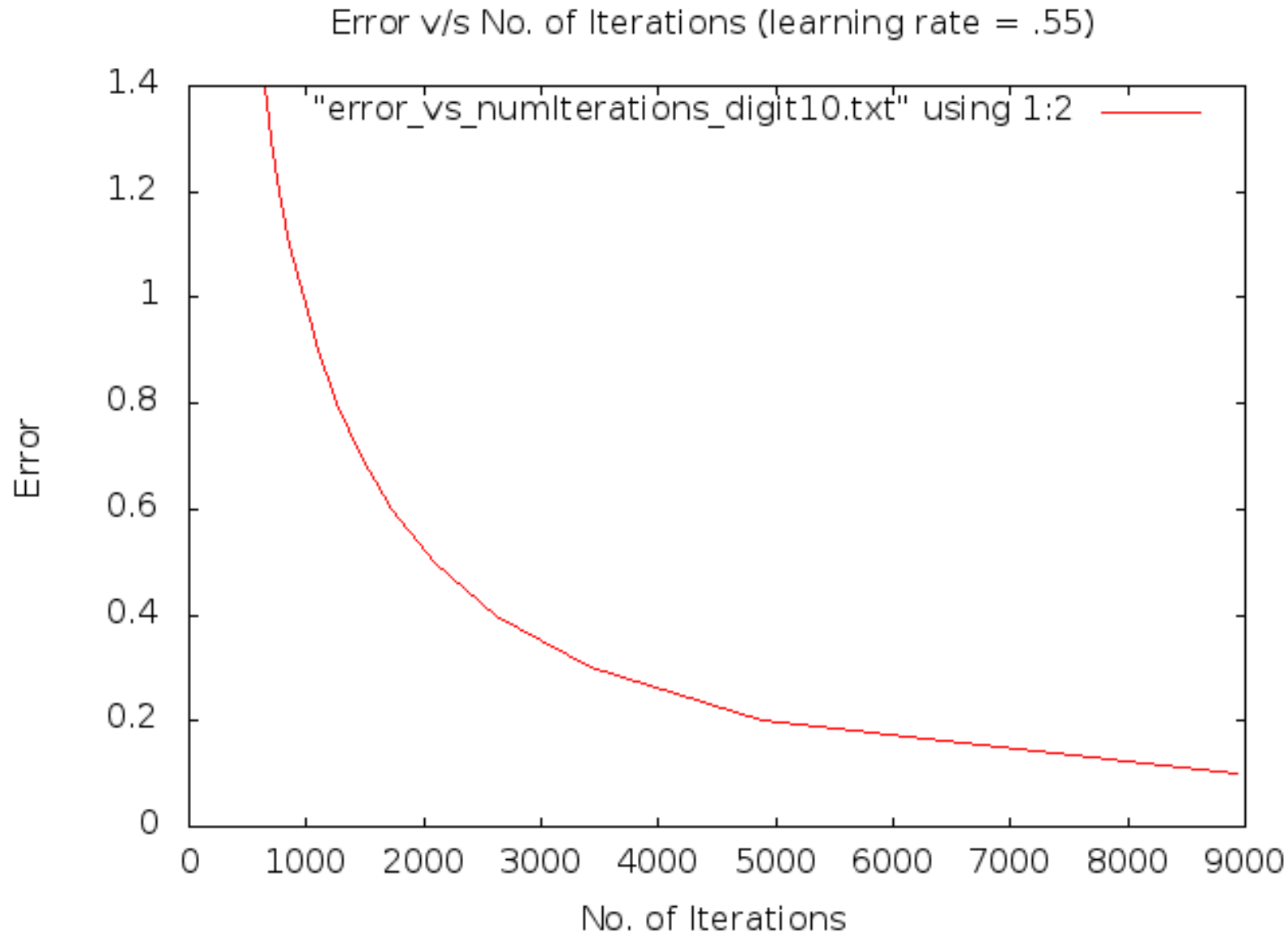
# Effect of Learning Rate (Digit Recogniser)



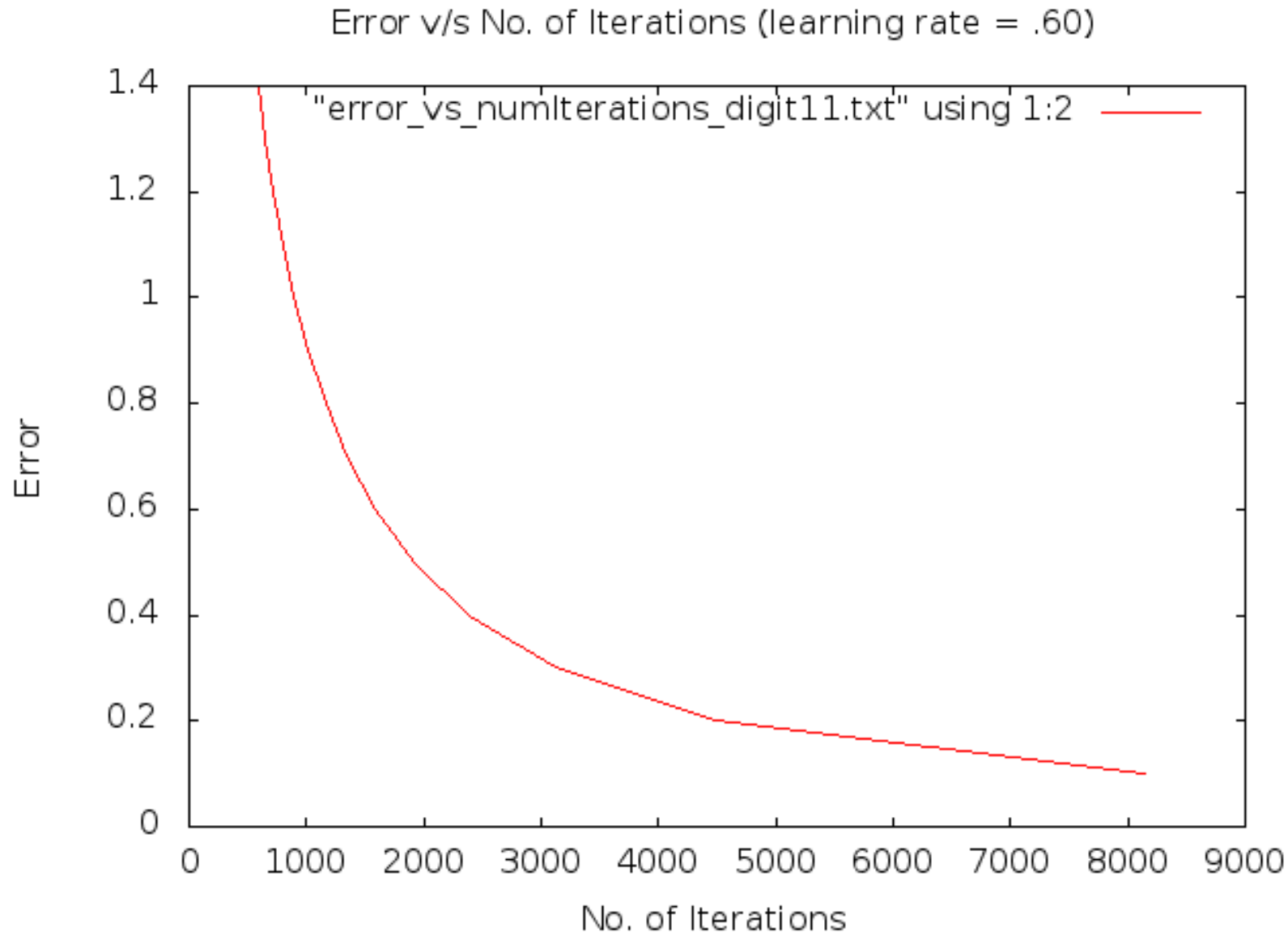
# Effect of Learning Rate (Digit Recogniser)



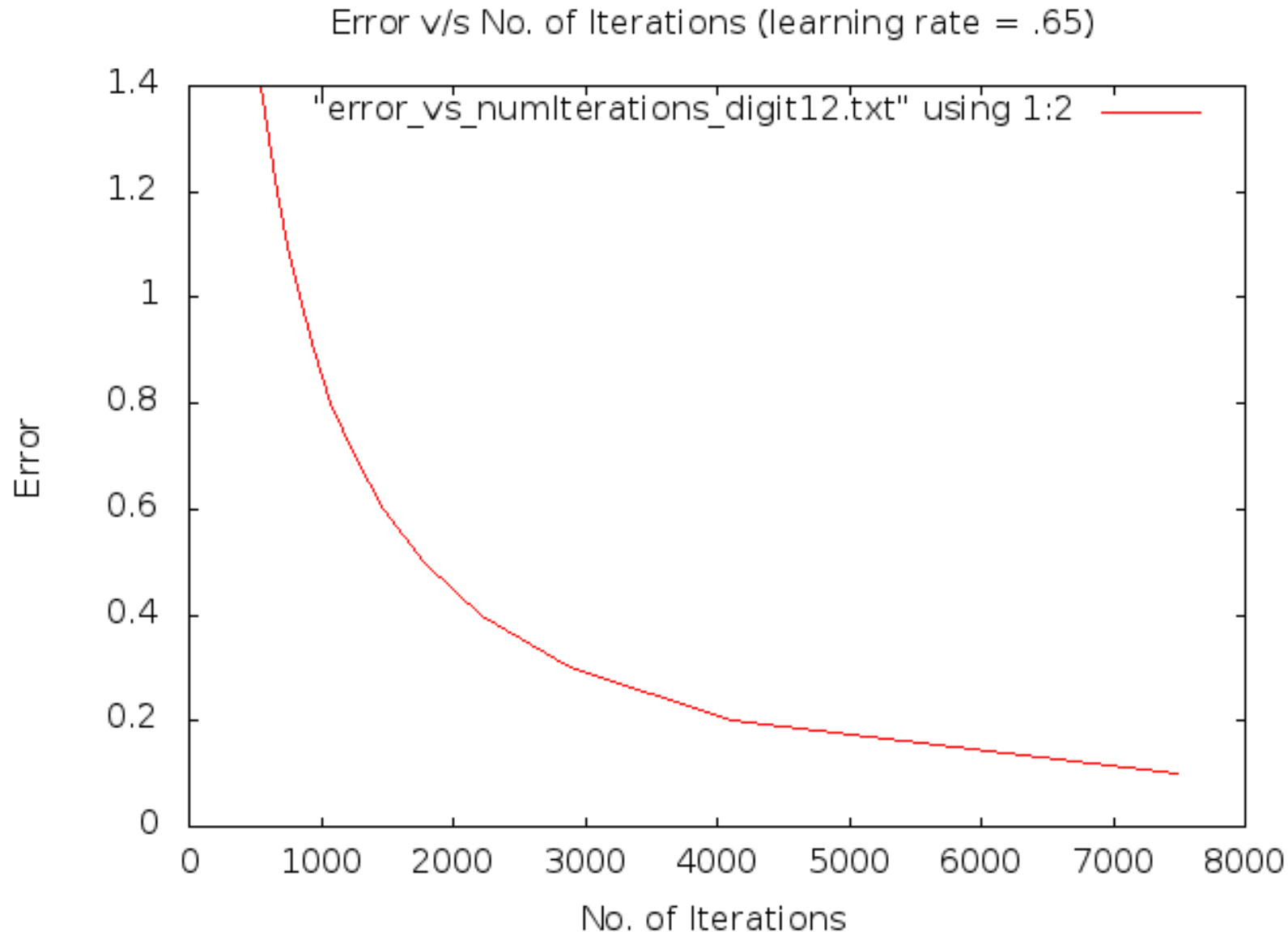
# Effect of Learning Rate (Digit Recogniser)



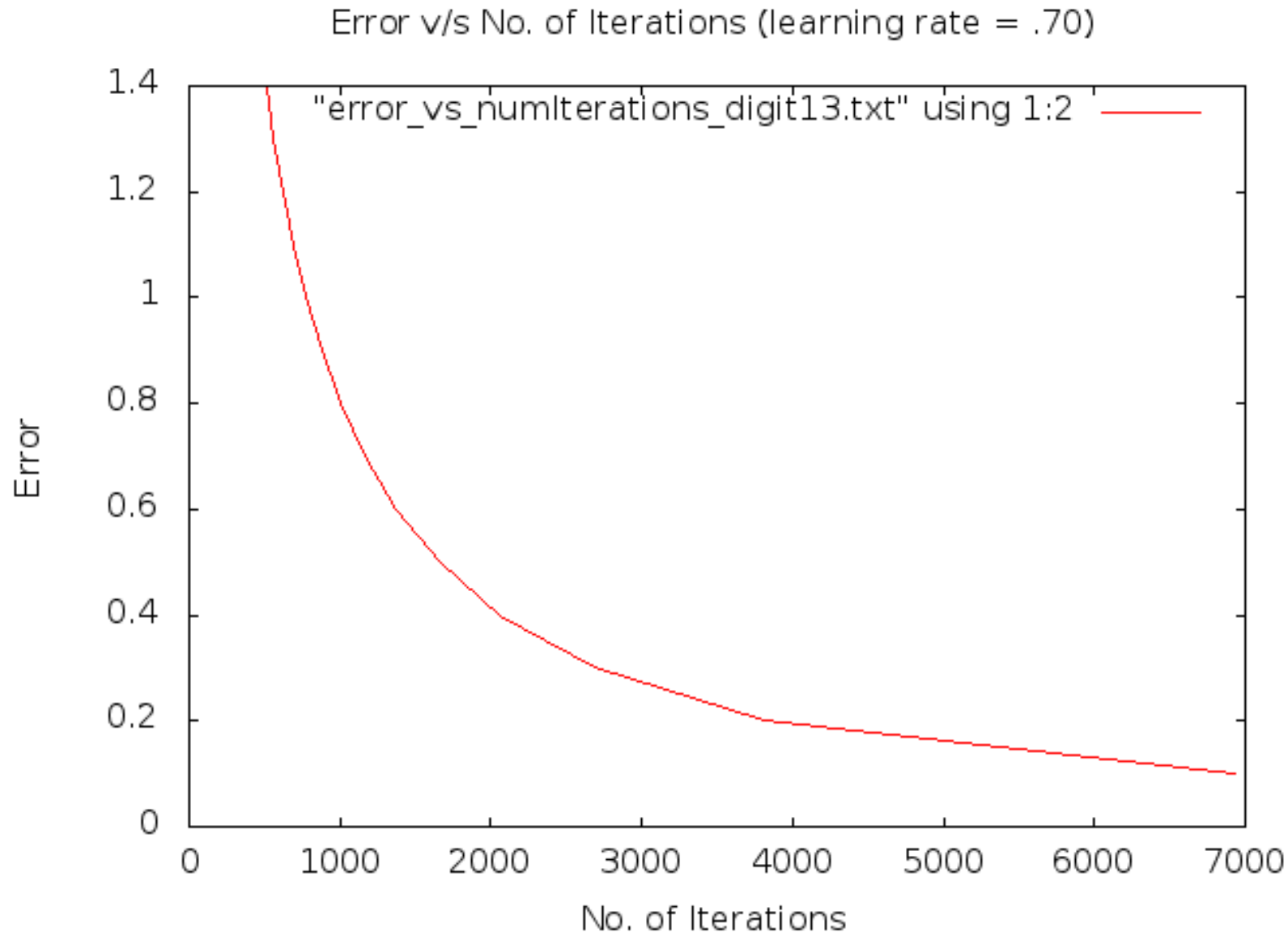
# Effect of Learning Rate (Digit Recogniser)



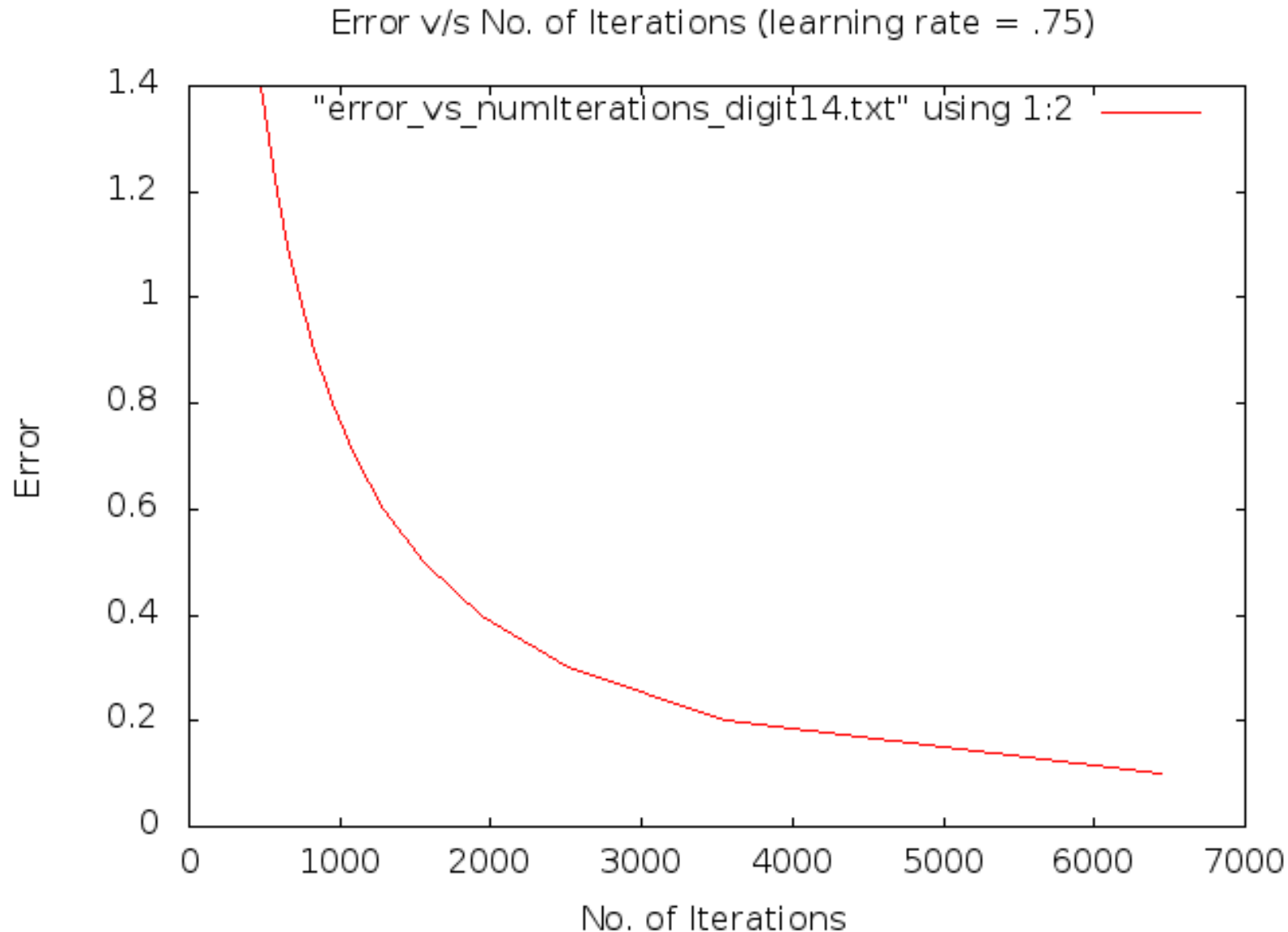
# Effect of Learning Rate (Digit Recogniser)



# Effect of Learning Rate (Digit Recogniser)

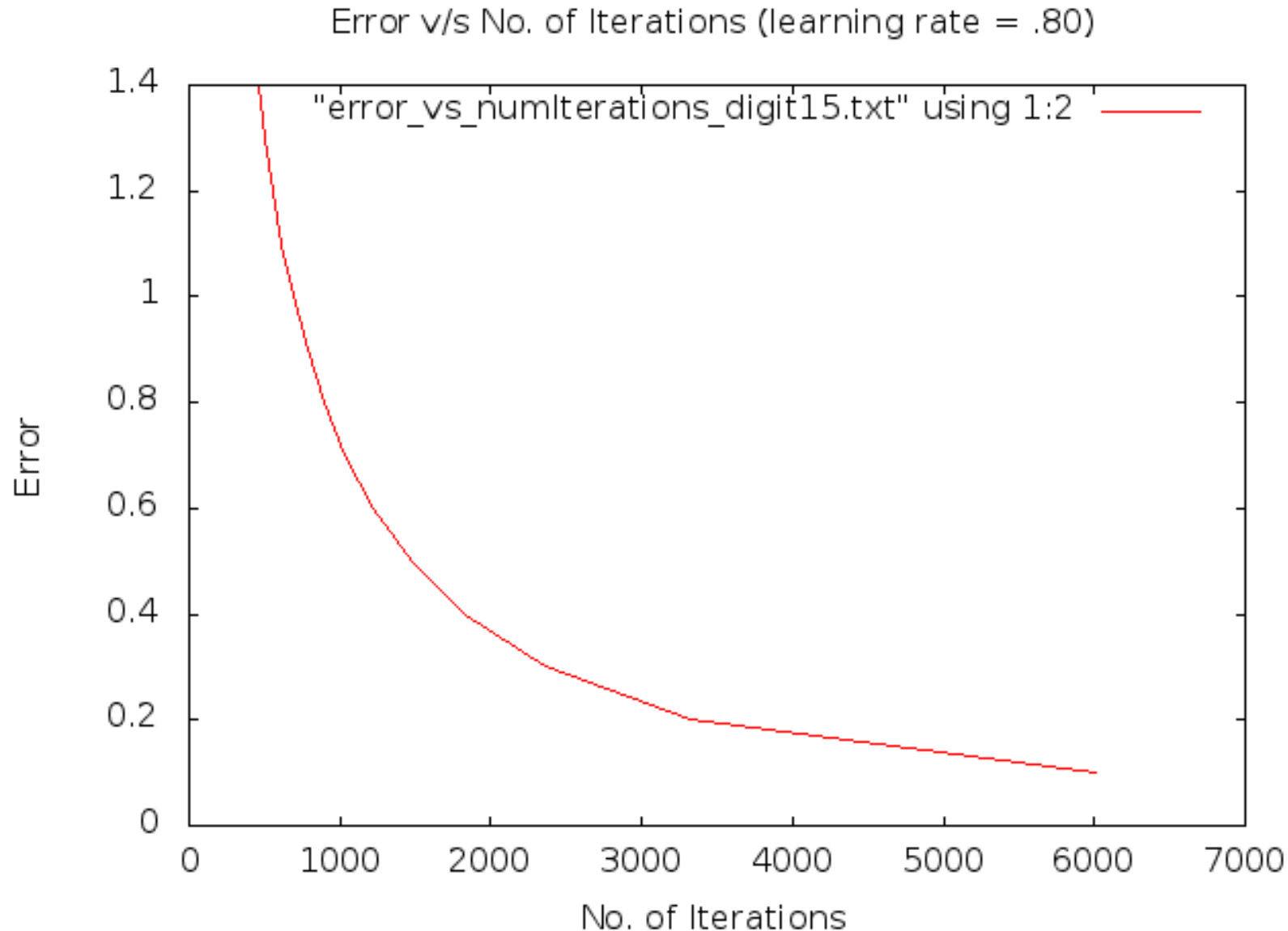


# Effect of Learning Rate (Digit Recogniser)

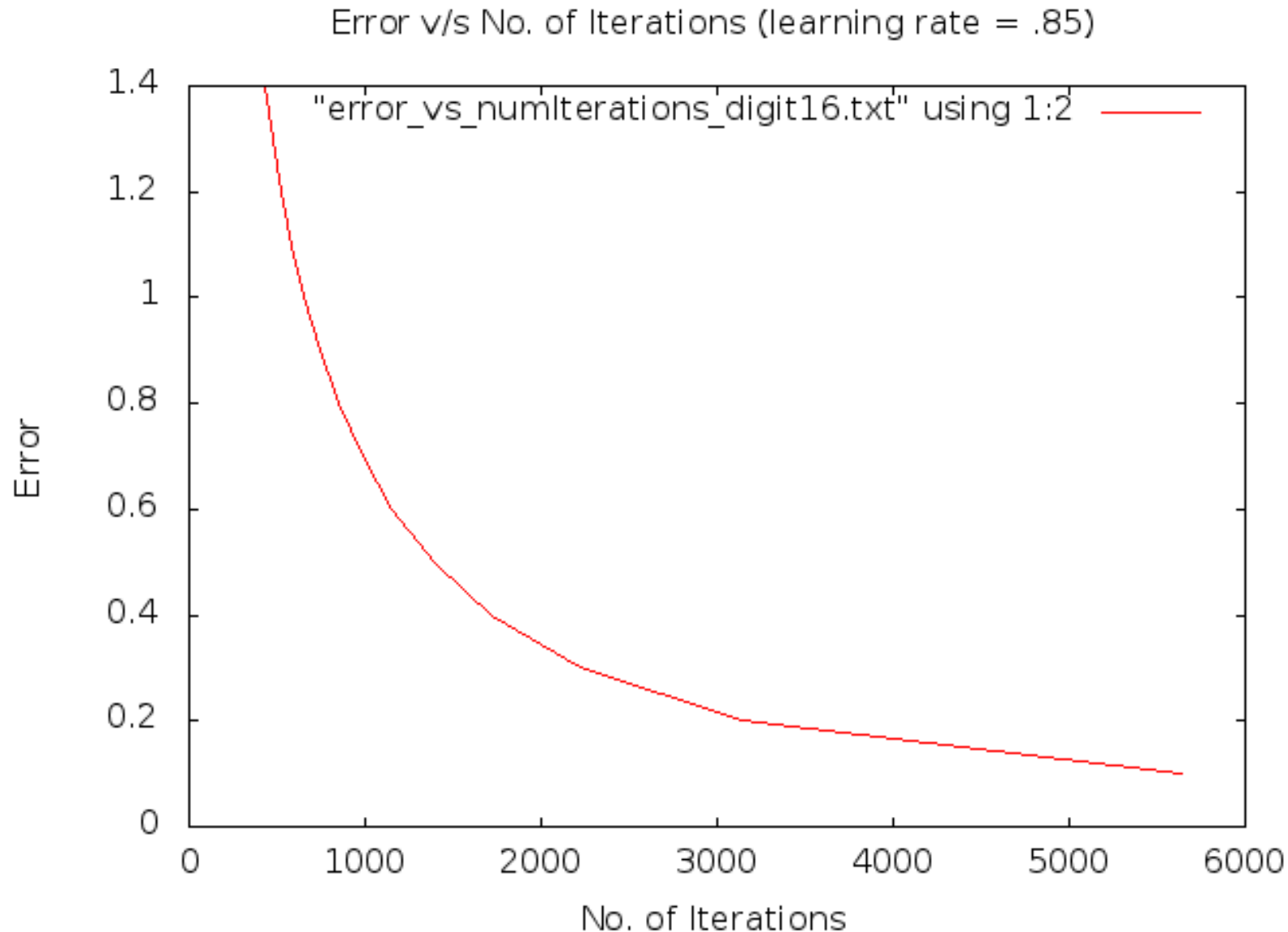




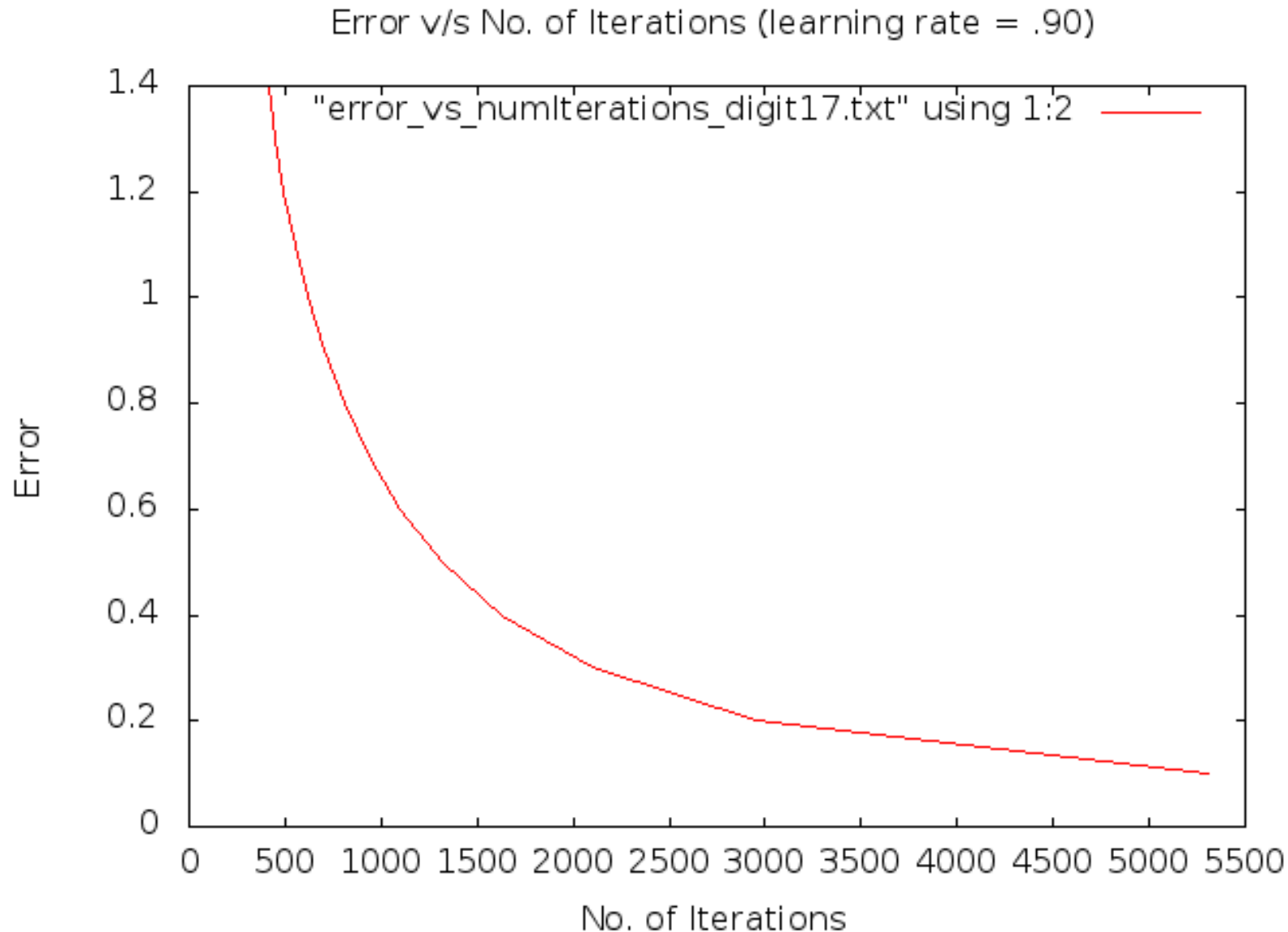
# Effect of Learning Rate (Digit Recogniser)



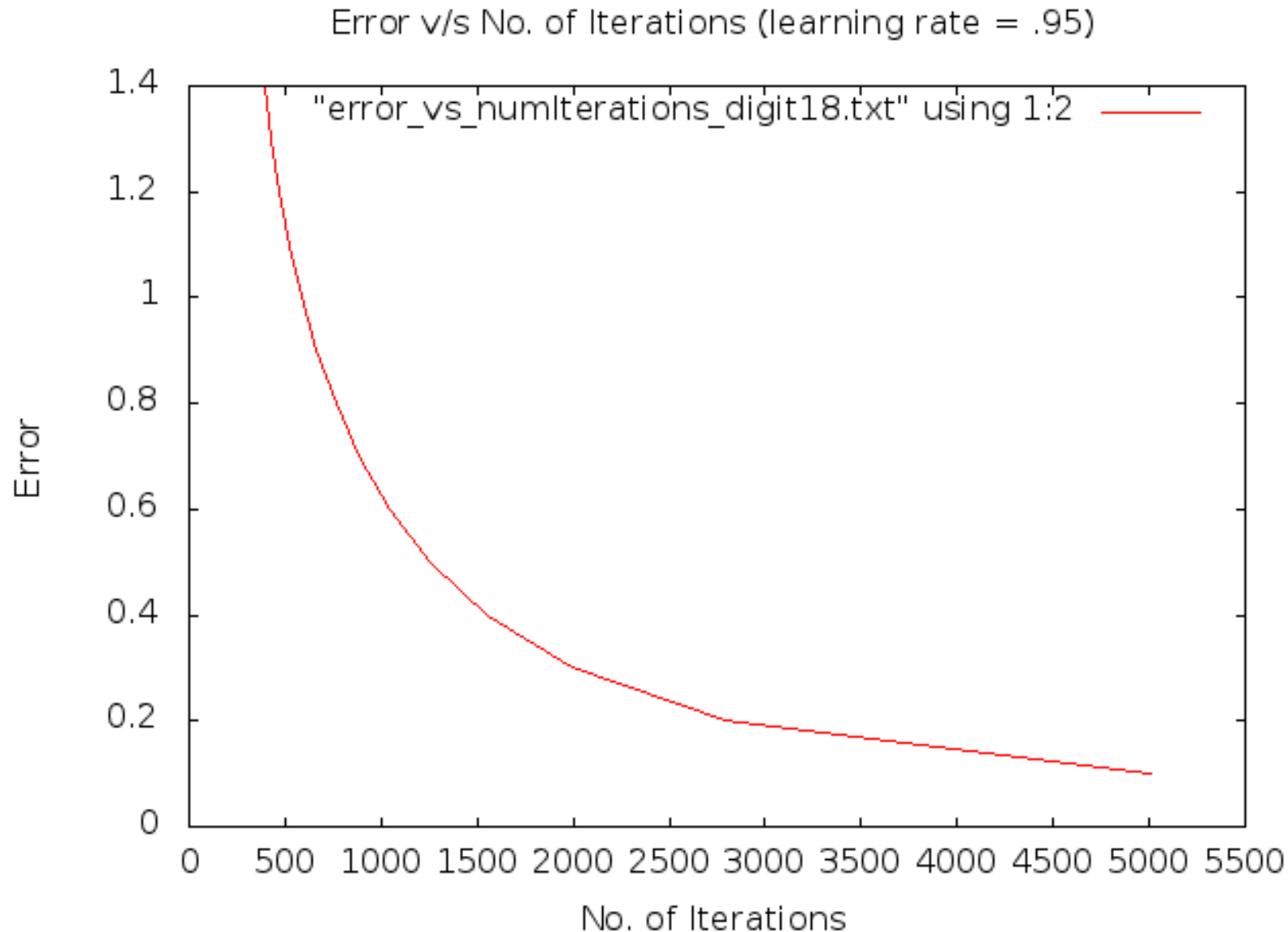
# Effect of Learning Rate (Digit Recogniser)



# Effect of Learning Rate (Digit Recogniser)

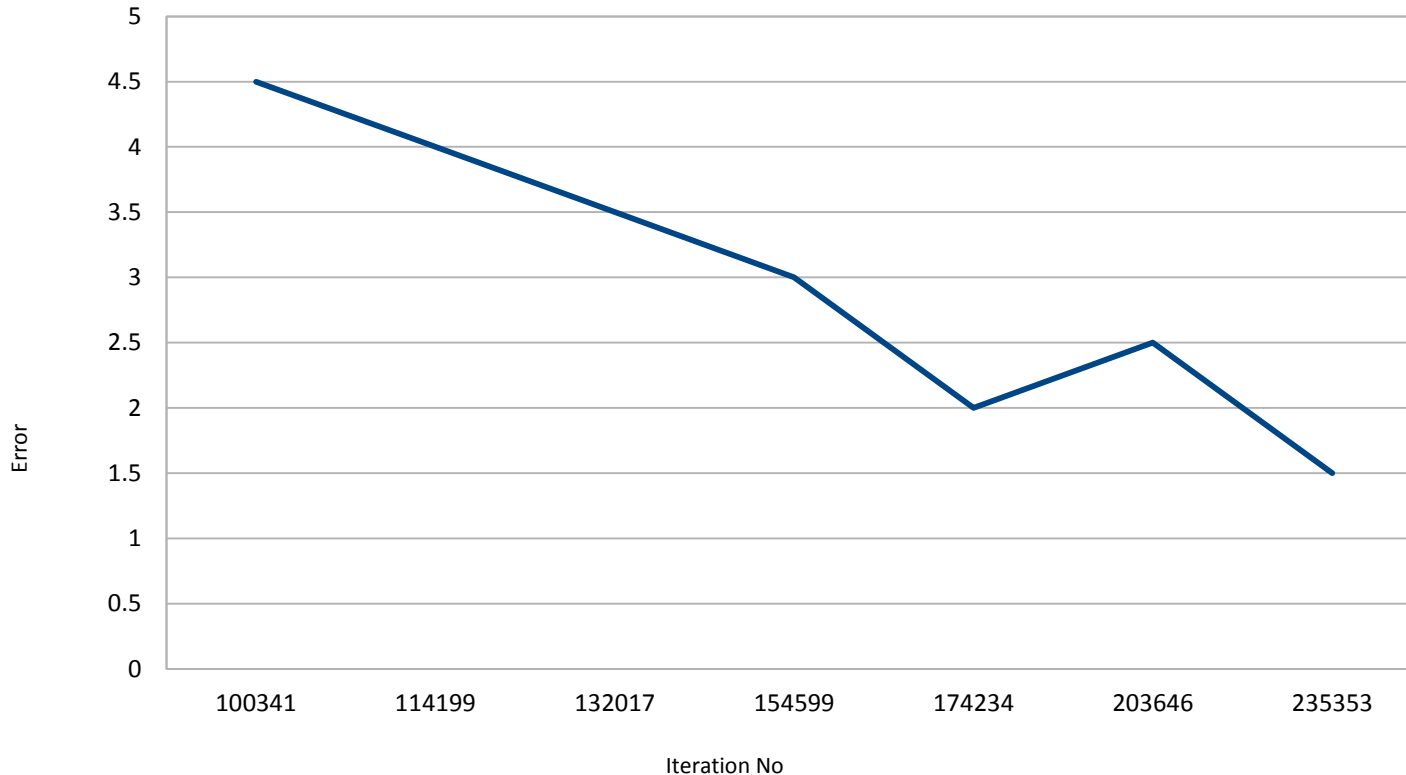


# Effect of Learning Rate (Digit Recogniser)



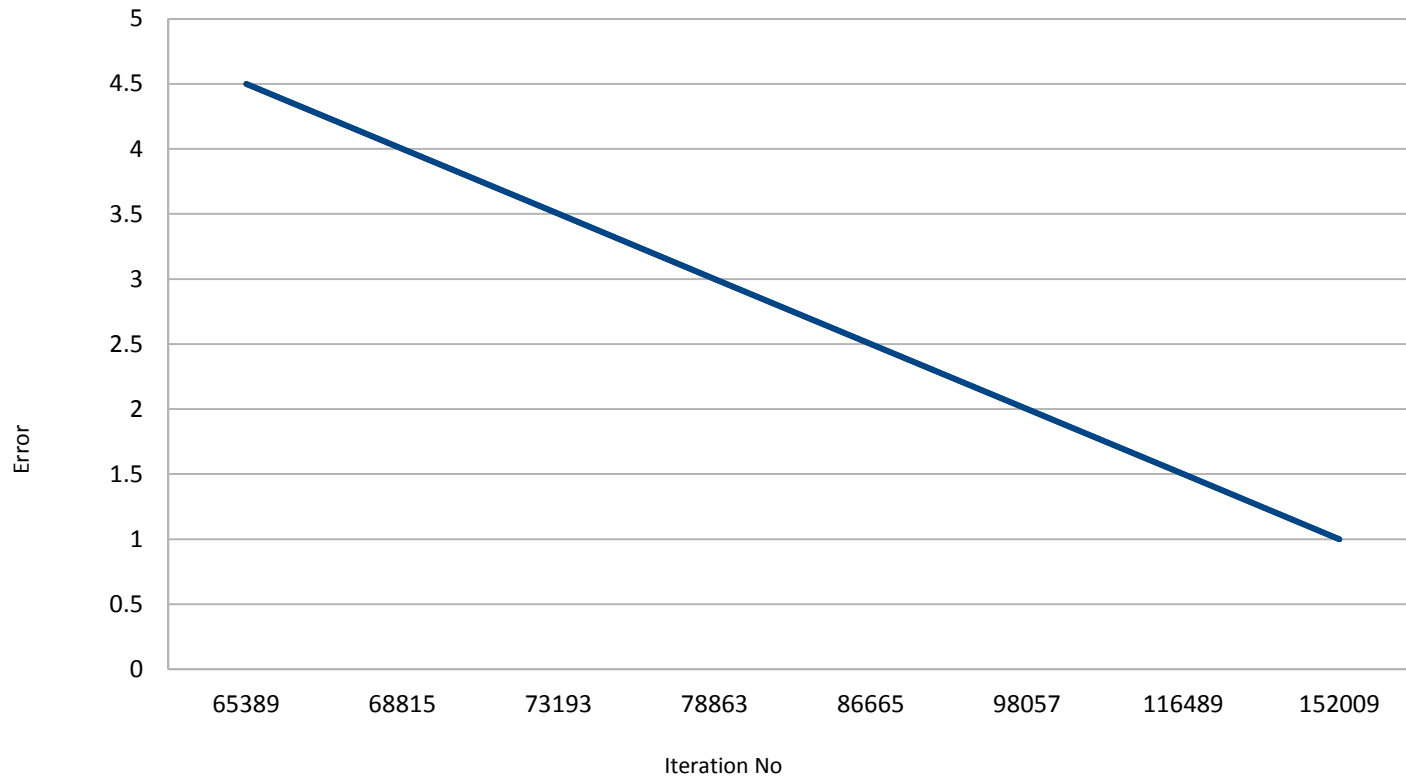
# Effect of Learning Rate (Parity)

Eta=0.1 Momentum Factor=0.05



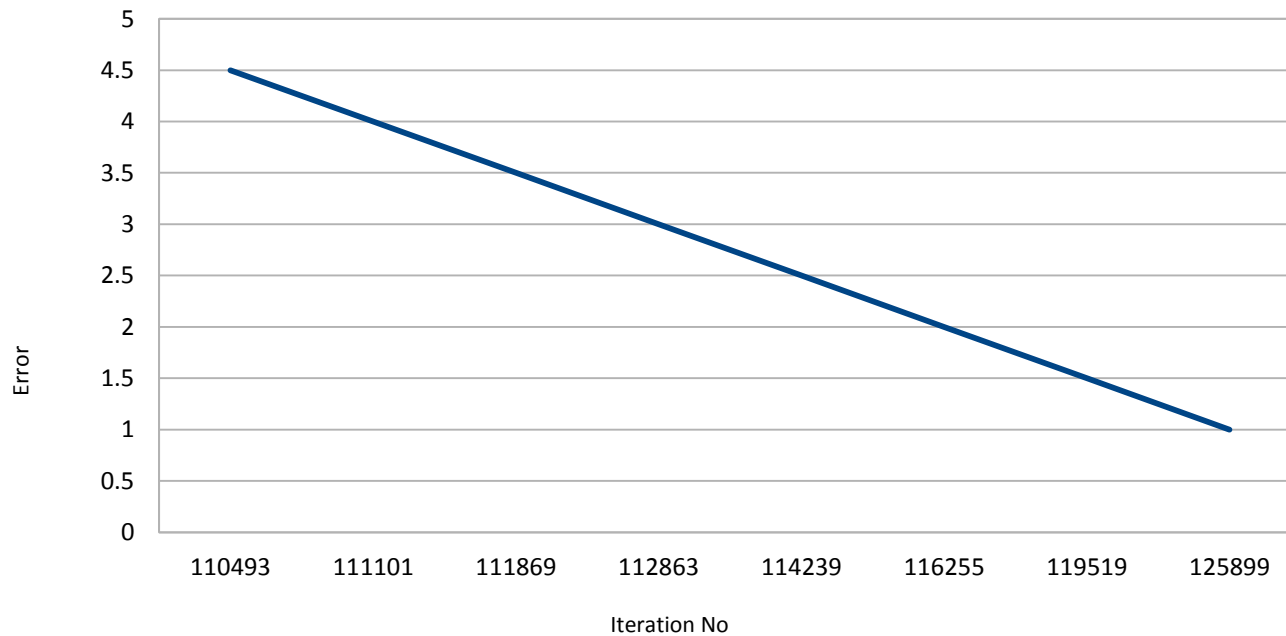
# Effect of Learning Rate (Parity)

Eta=0.3 Momentum Factor=0.05



# Effect of Learning Rate (Parity)

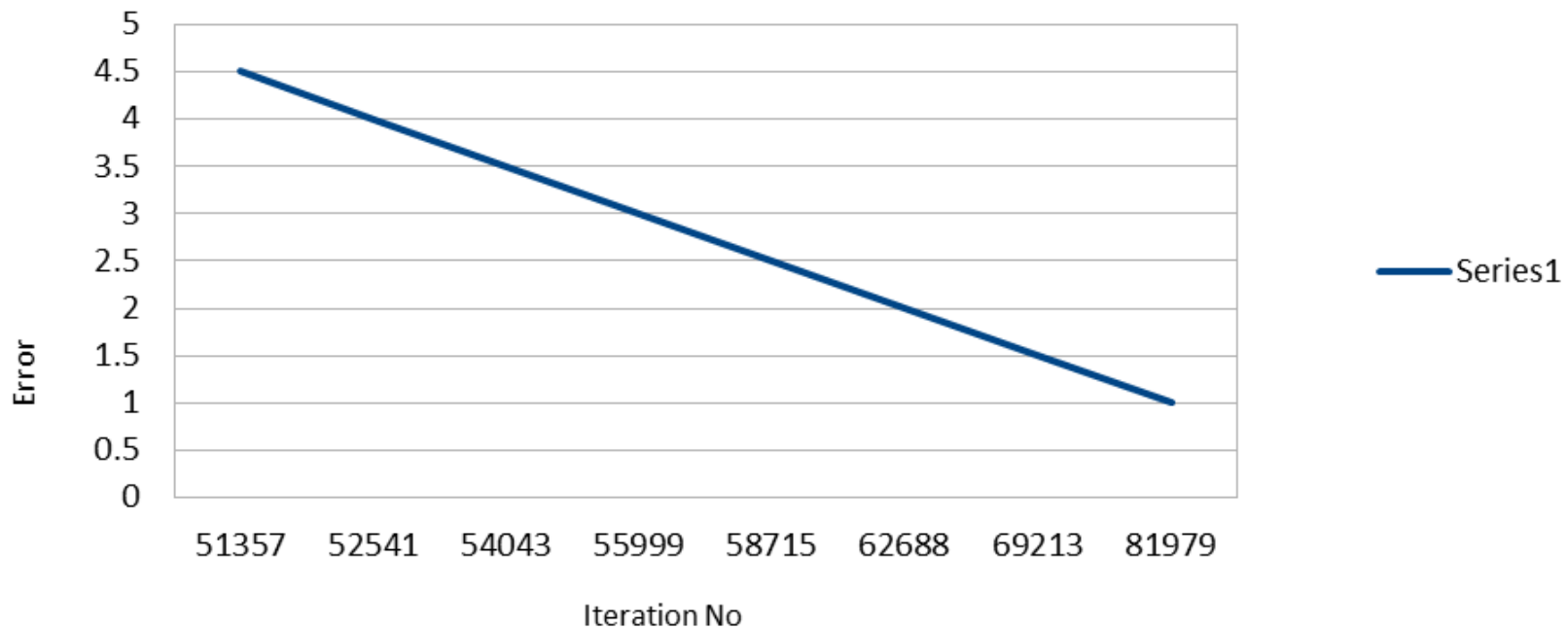
Eta=0.5 Momentum Factor=0.05



# Effect of Learning Rate (Parity)

Eta=0.7, Momentum Factor=0.05

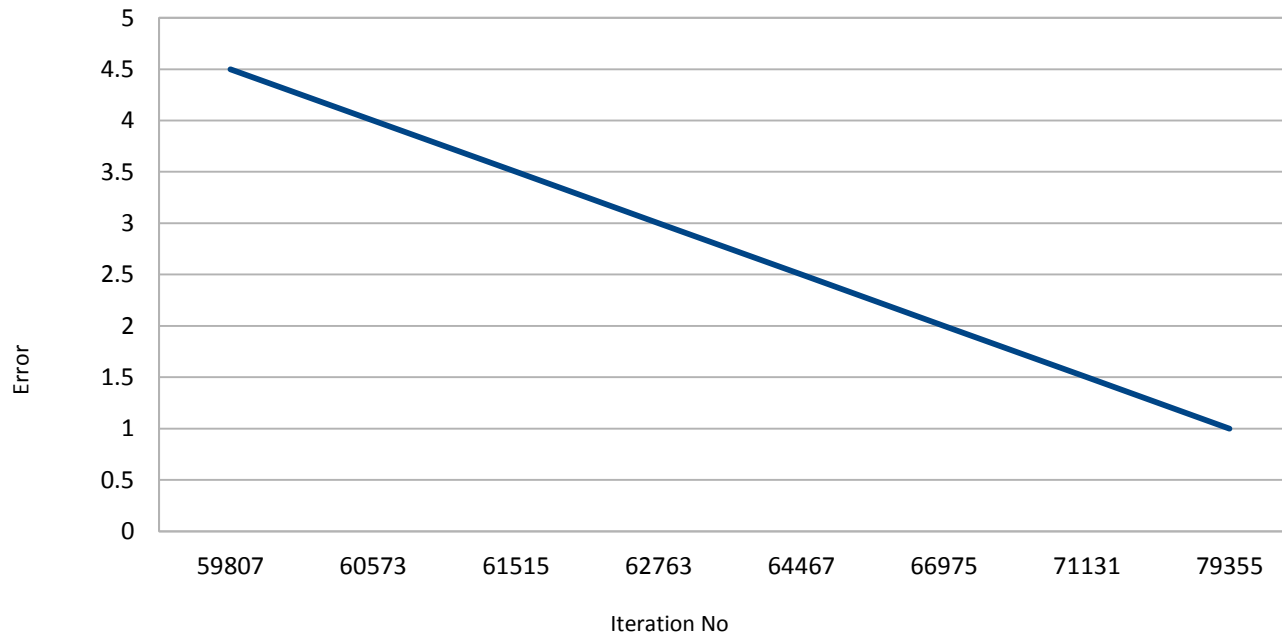
Error vs Iteration No





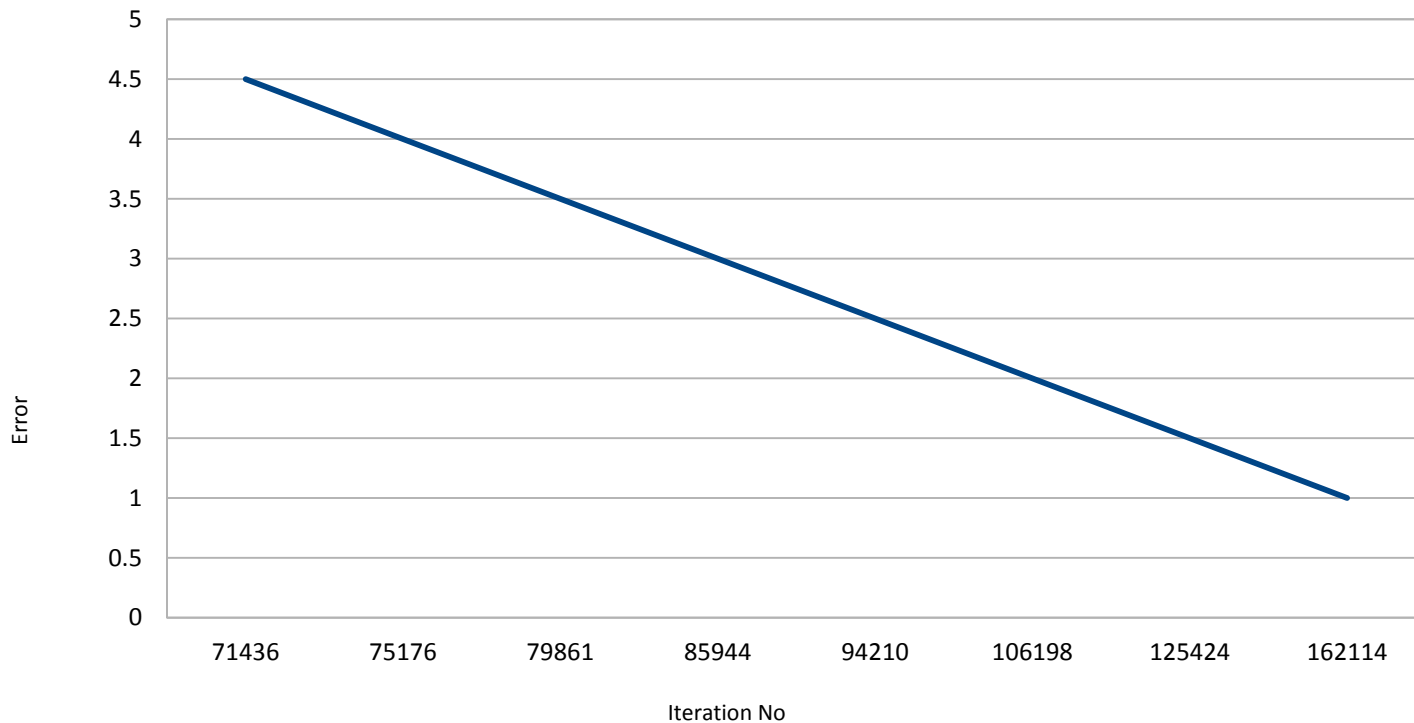
# Effect of Learning Rate (Parity)

Eta=0.9, Momentum Factor=0.05



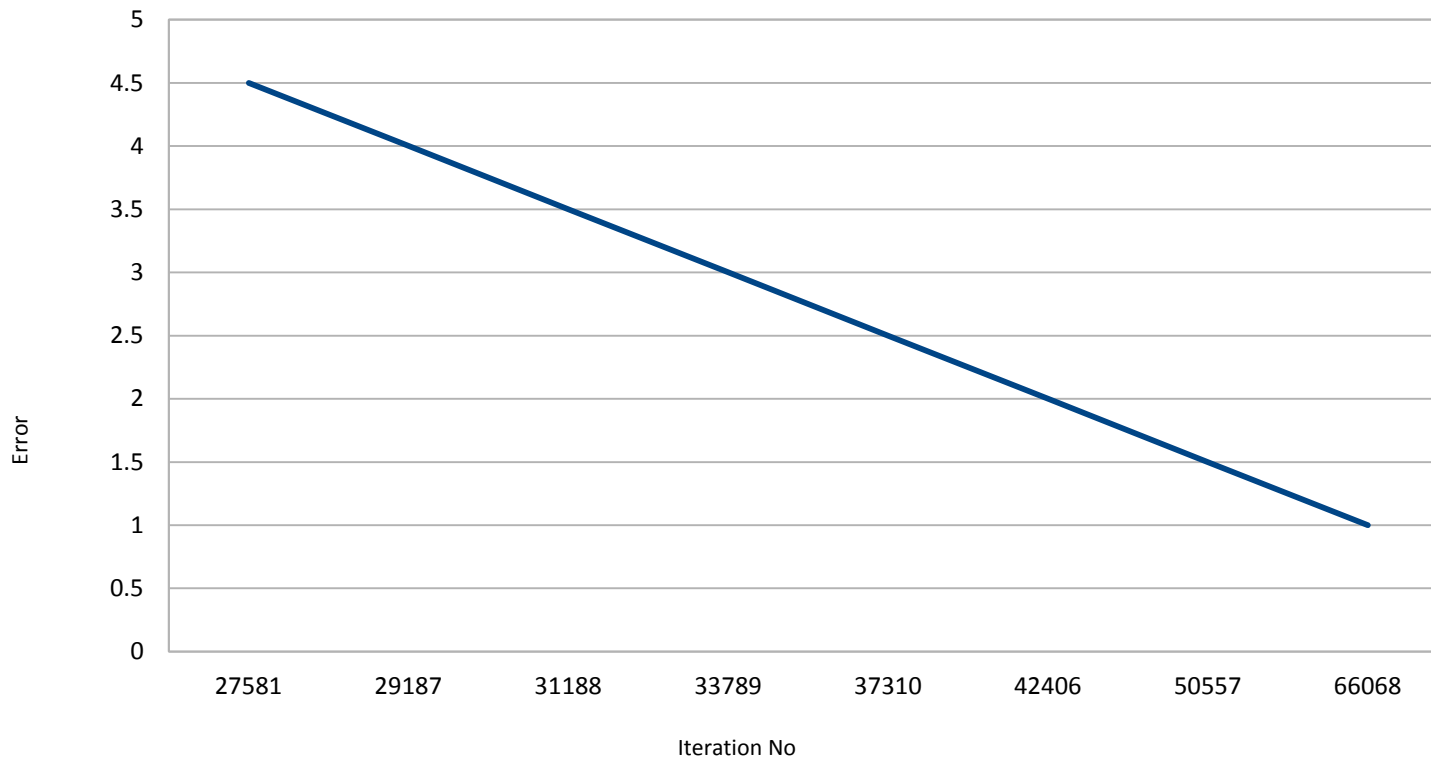
# Effect of Learning Rate (Palindrome)

Eta=0.1 Momentum Factor=0.05



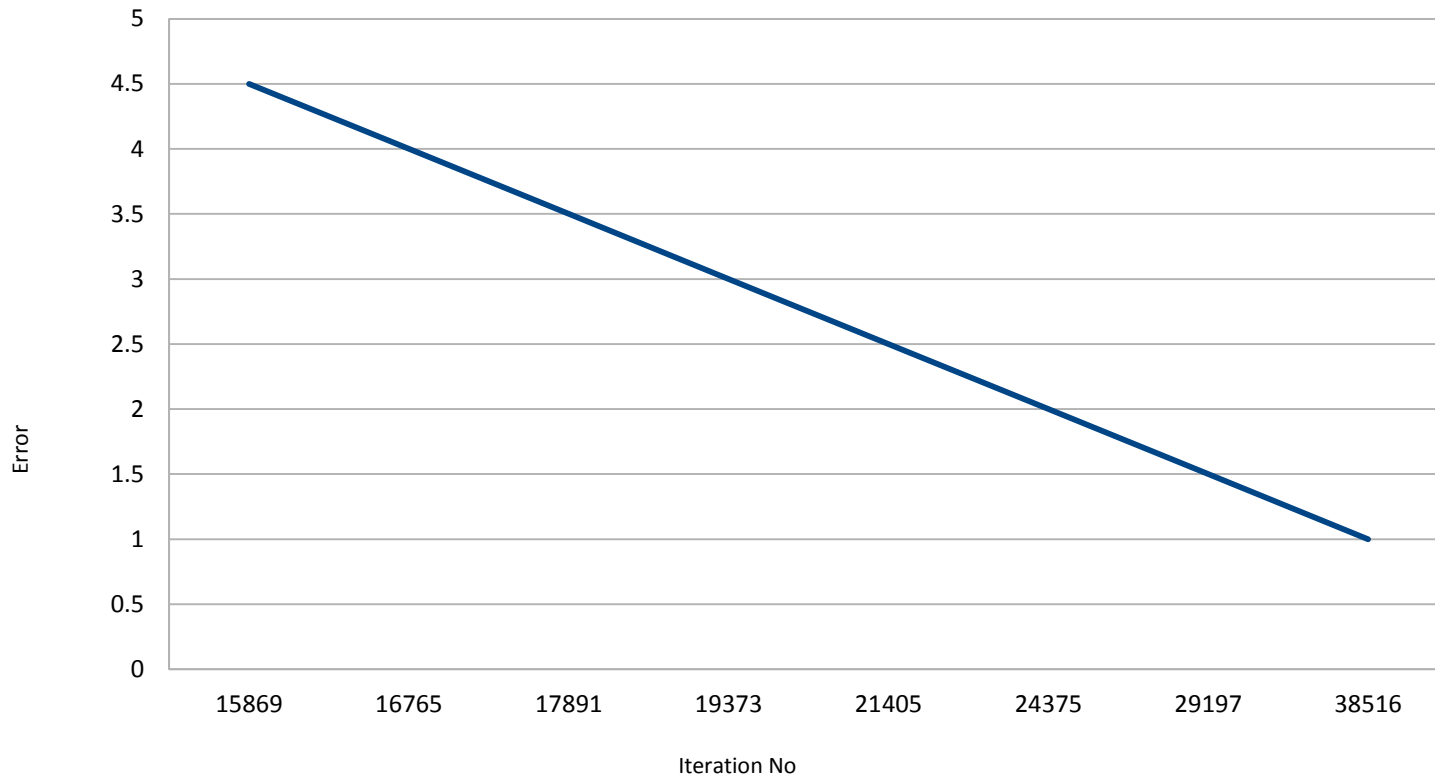
# Effect of Learning Rate (Palindrome)

Eta=0.3 Momentum Factor=0.05



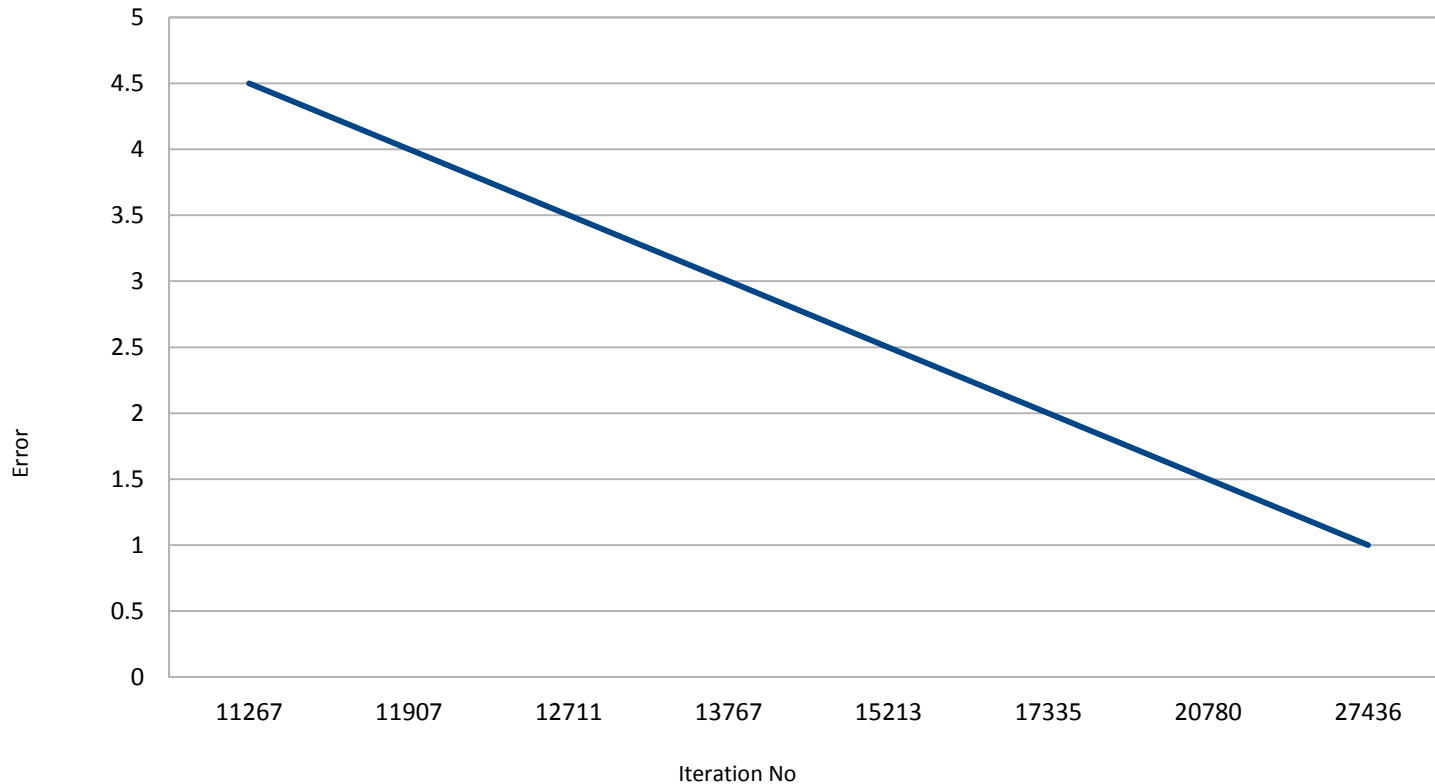
# Effect of Learning Rate (Palindrome)

Eta=0.5 Momentum Factor=0.05



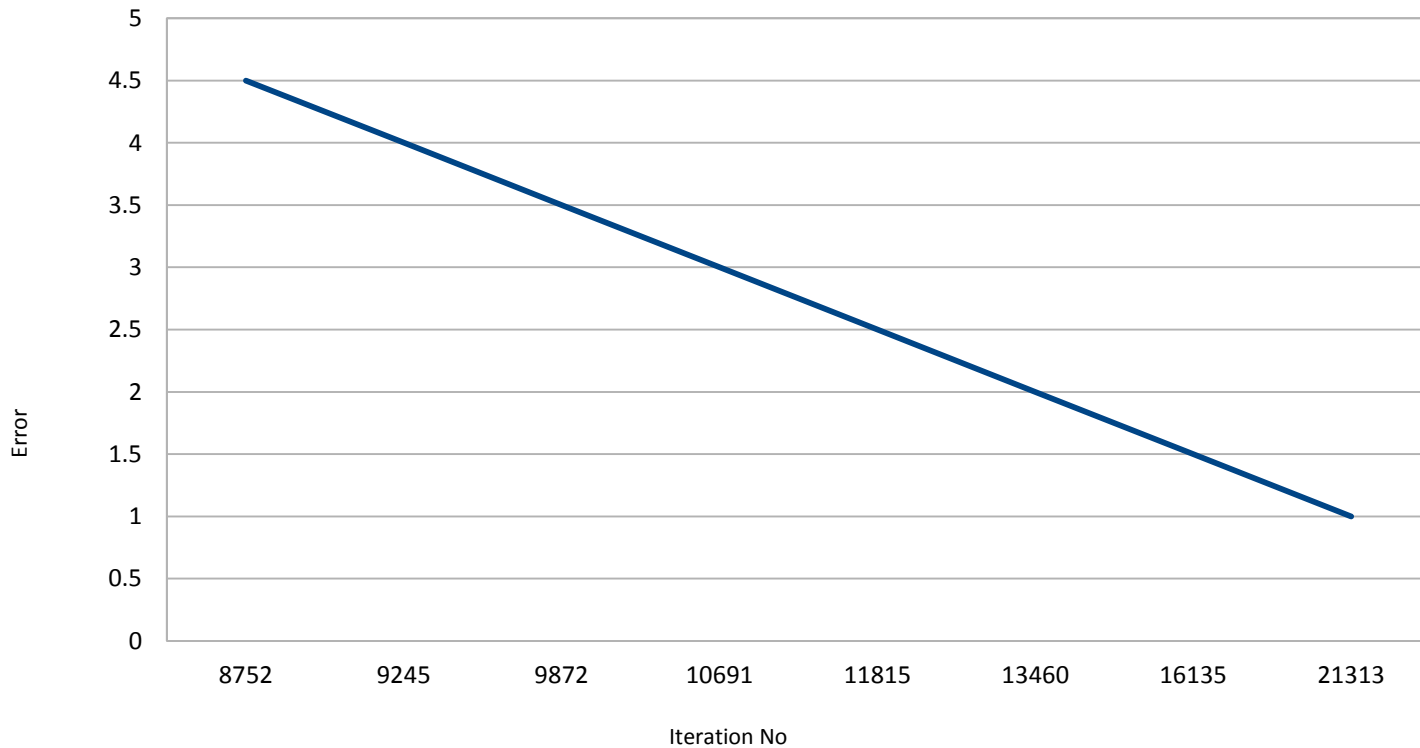
# Effect of Learning Rate (Palindrome)

Eta=0.7 Momentum Factor=0.05



# Effect of Learning Rate (Palindrome)

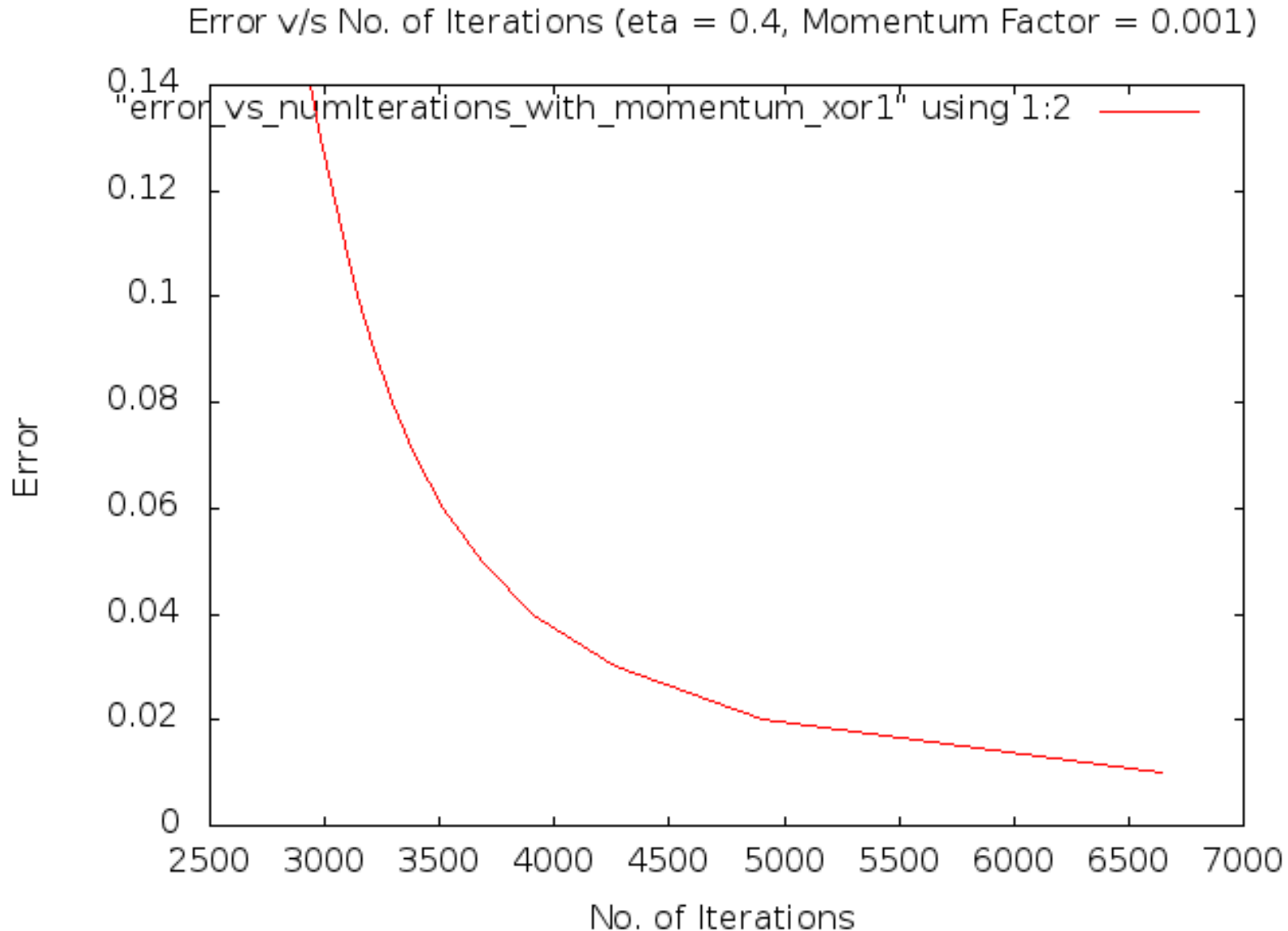
Eta=0.9 Momentum Factor=0.05



# Effect of Learning Rate

- We varied the learning rate from 0.10 to 0.95 in intervals of 0.05 and observed the effect of learning rate.
- In general, as the learning rate increases, no of iterations required to converge to a solution decreases.
- The speed of convergence of a network can be improved by increasing the learning rate.
- Unfortunately, increasing learning rate usually results in increasing network instability, with weight values oscillating erratically as they converge on a solution.
- Higher values of learning rate may provide faster convergence on a solution, but may also increase instability and may lead to a failure to converge.

# Effect of Momentum Factor (XOR)

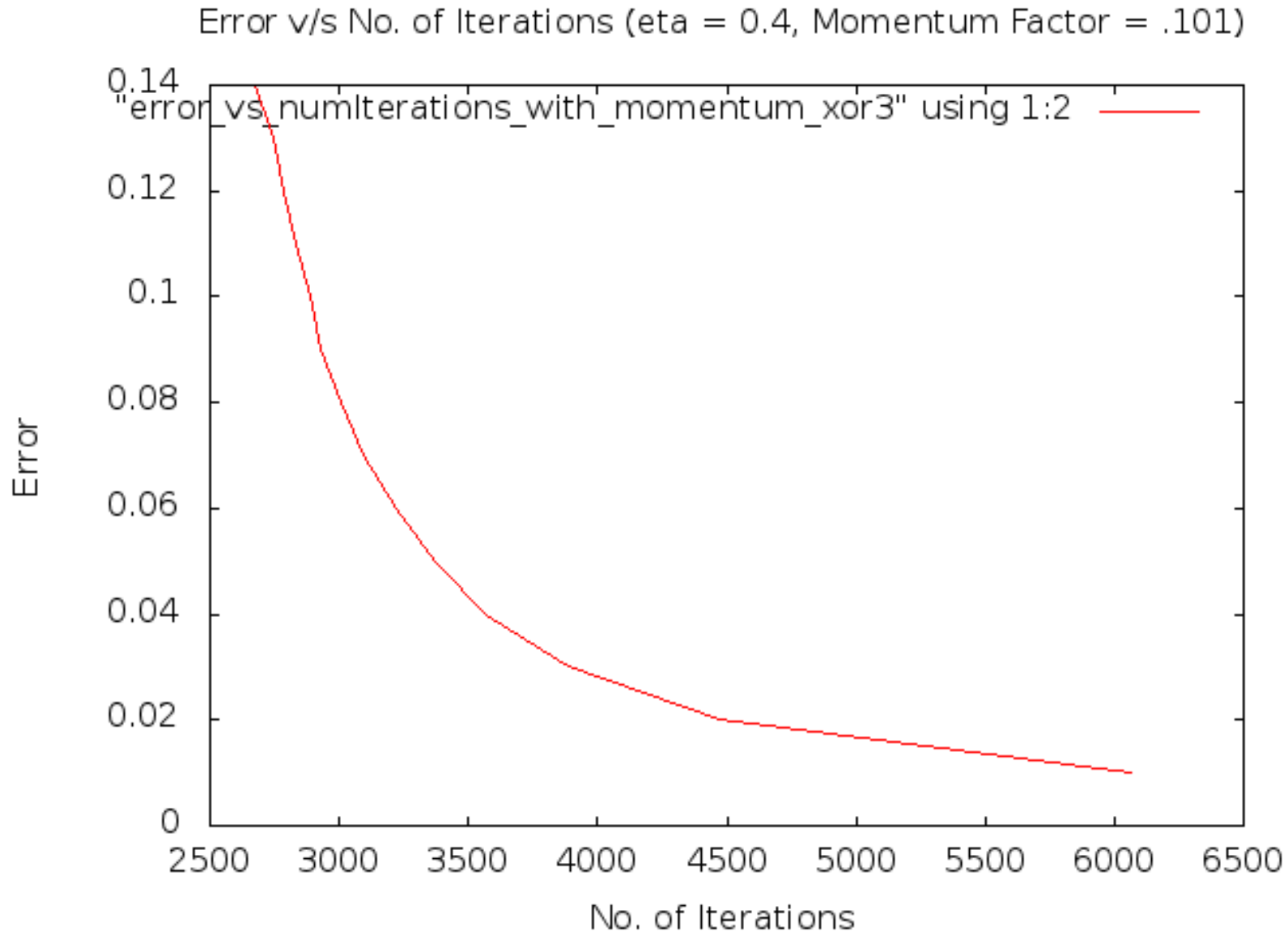




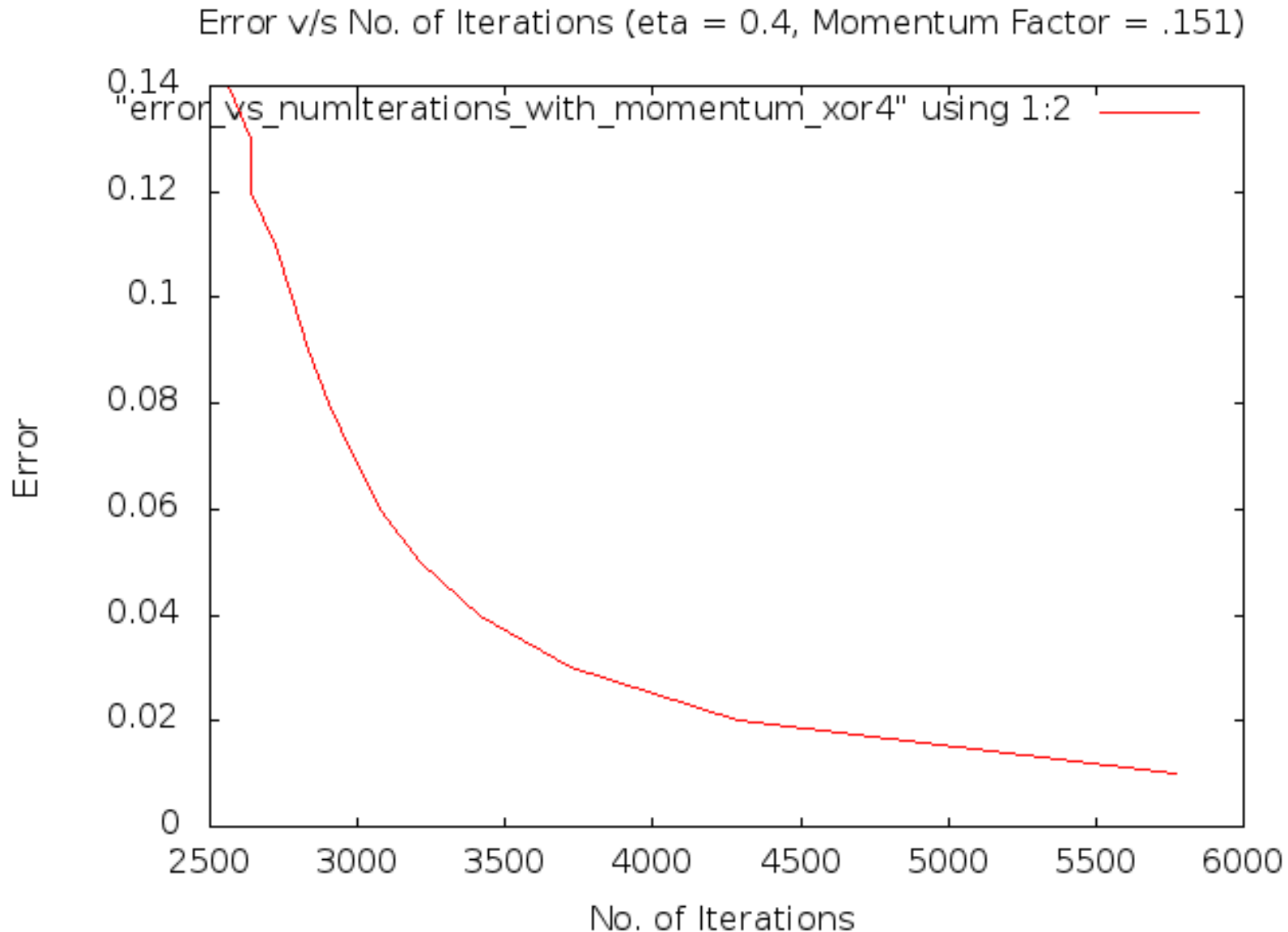
# Effect of Momentum Factor (XOR)



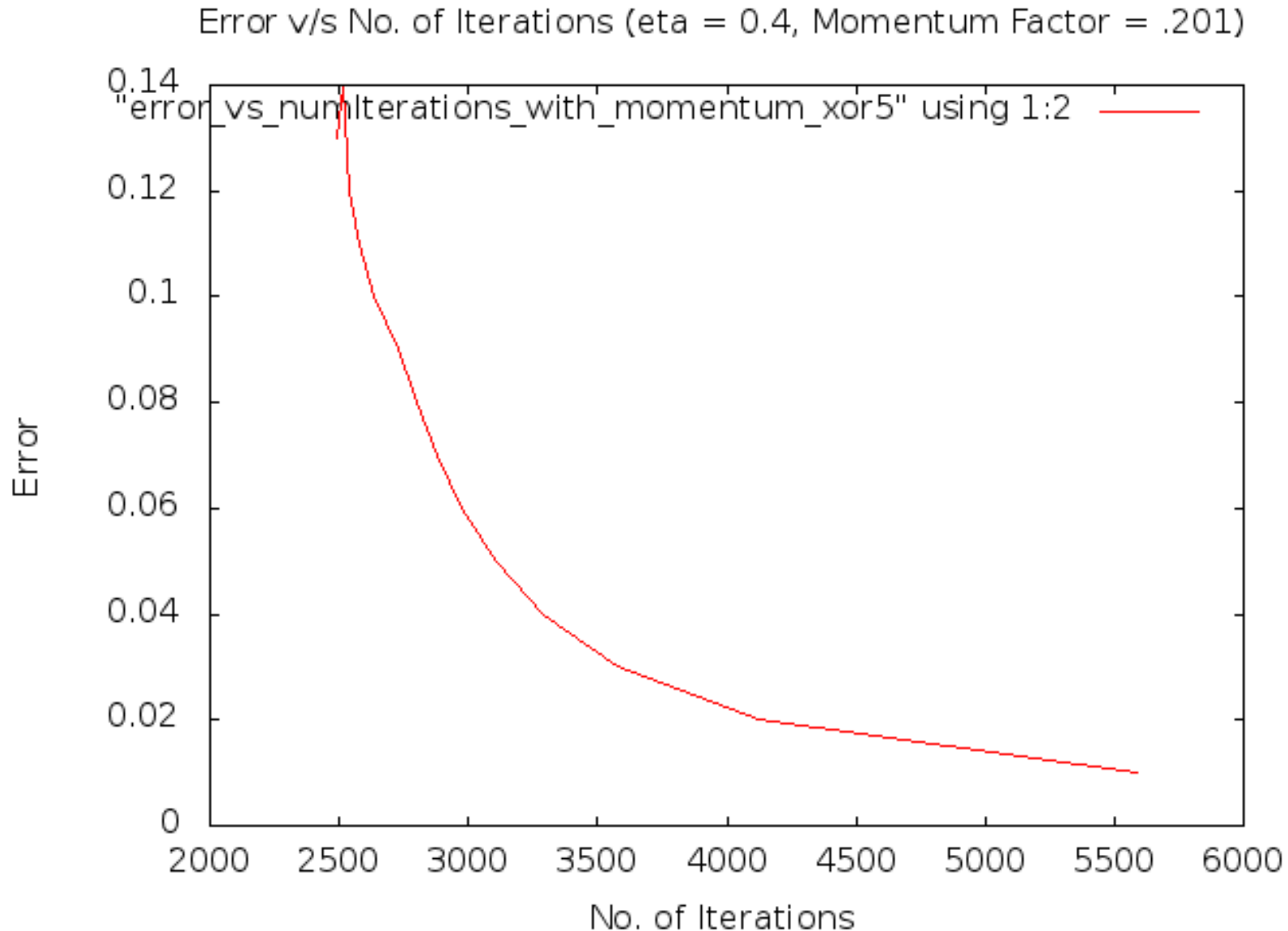
# Effect of Momentum Factor (XOR)



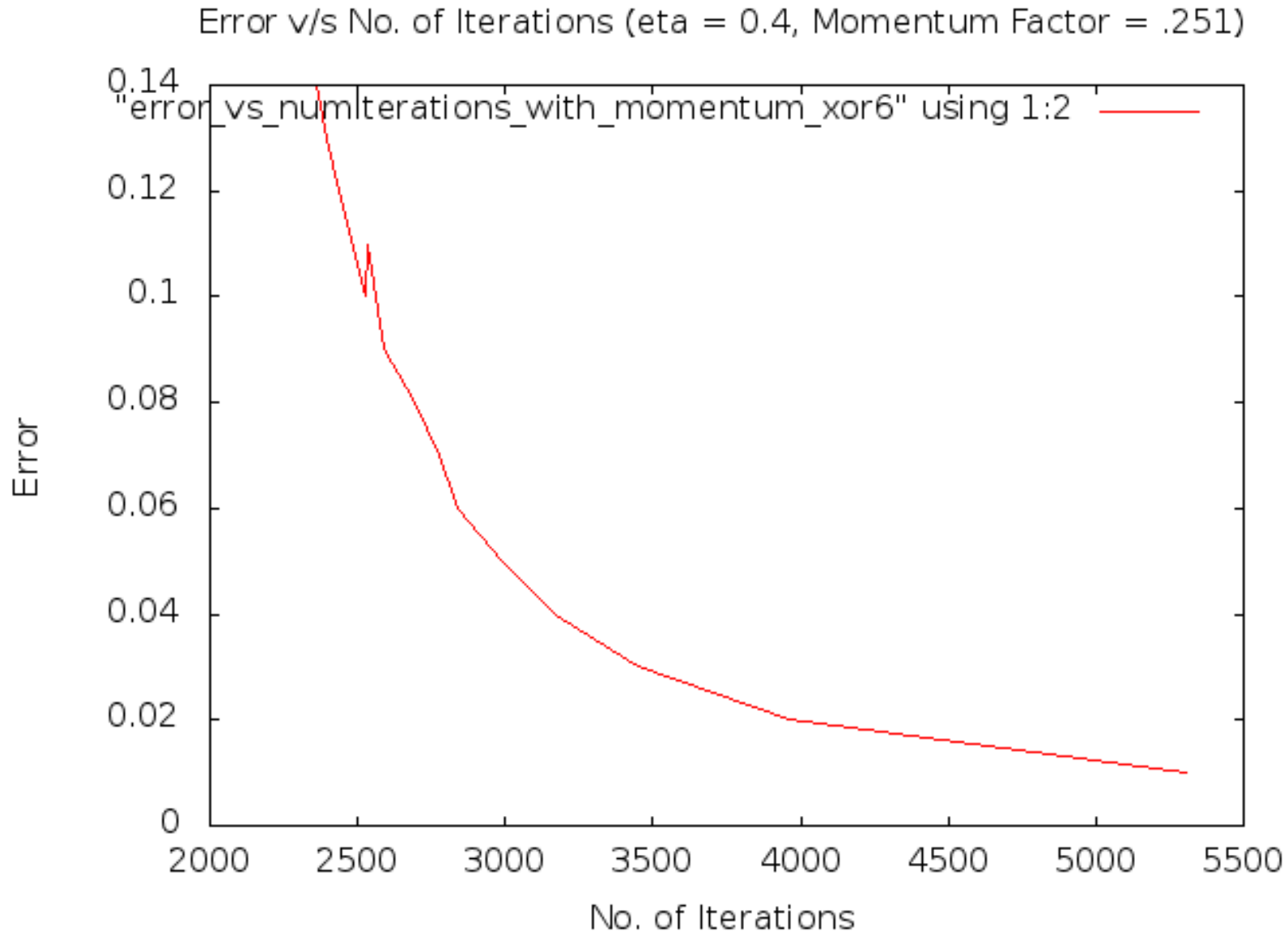
# Effect of Momentum Factor (XOR)



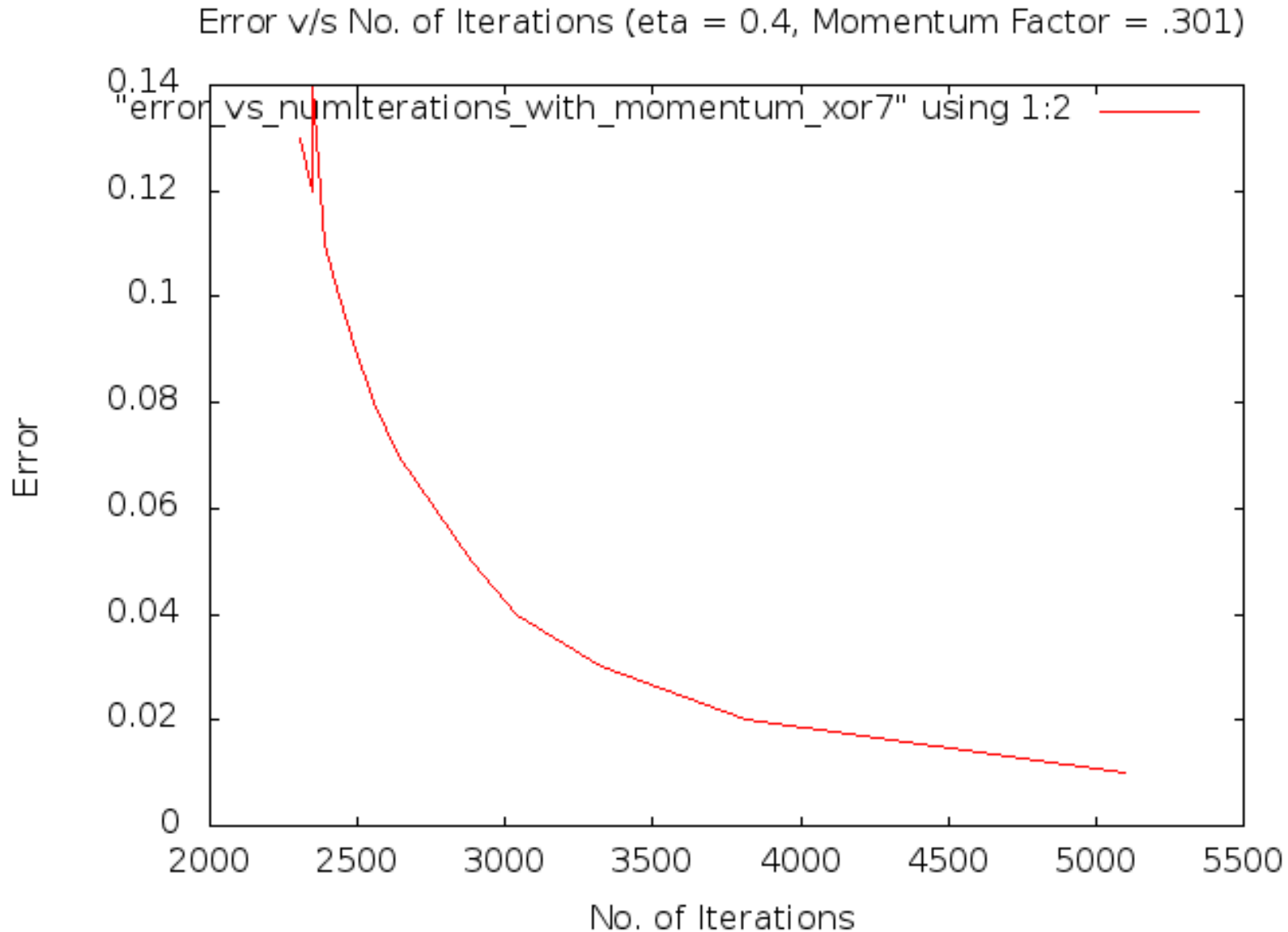
# Effect of Momentum Factor (XOR)



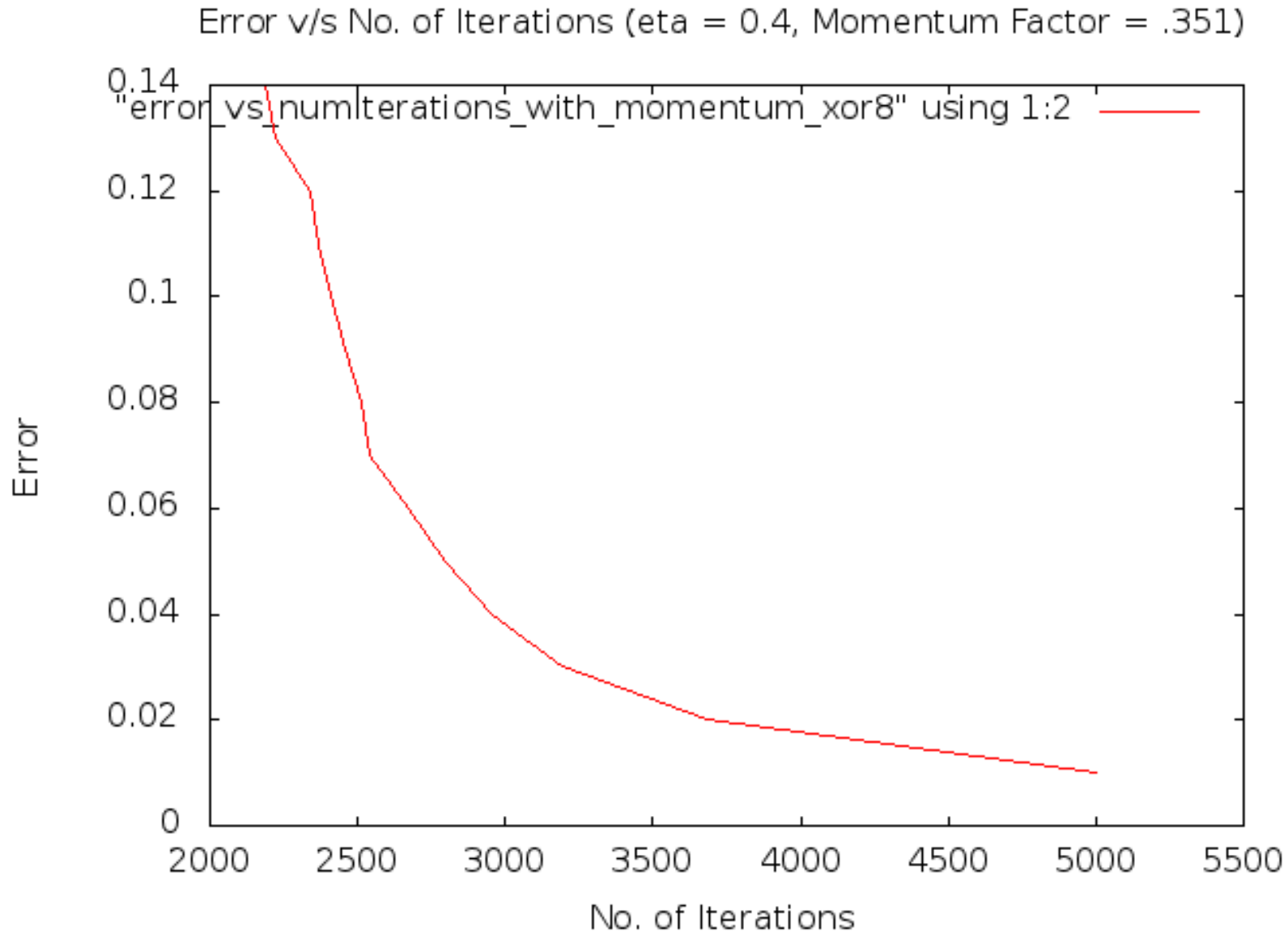
# Effect of Momentum Factor (XOR)



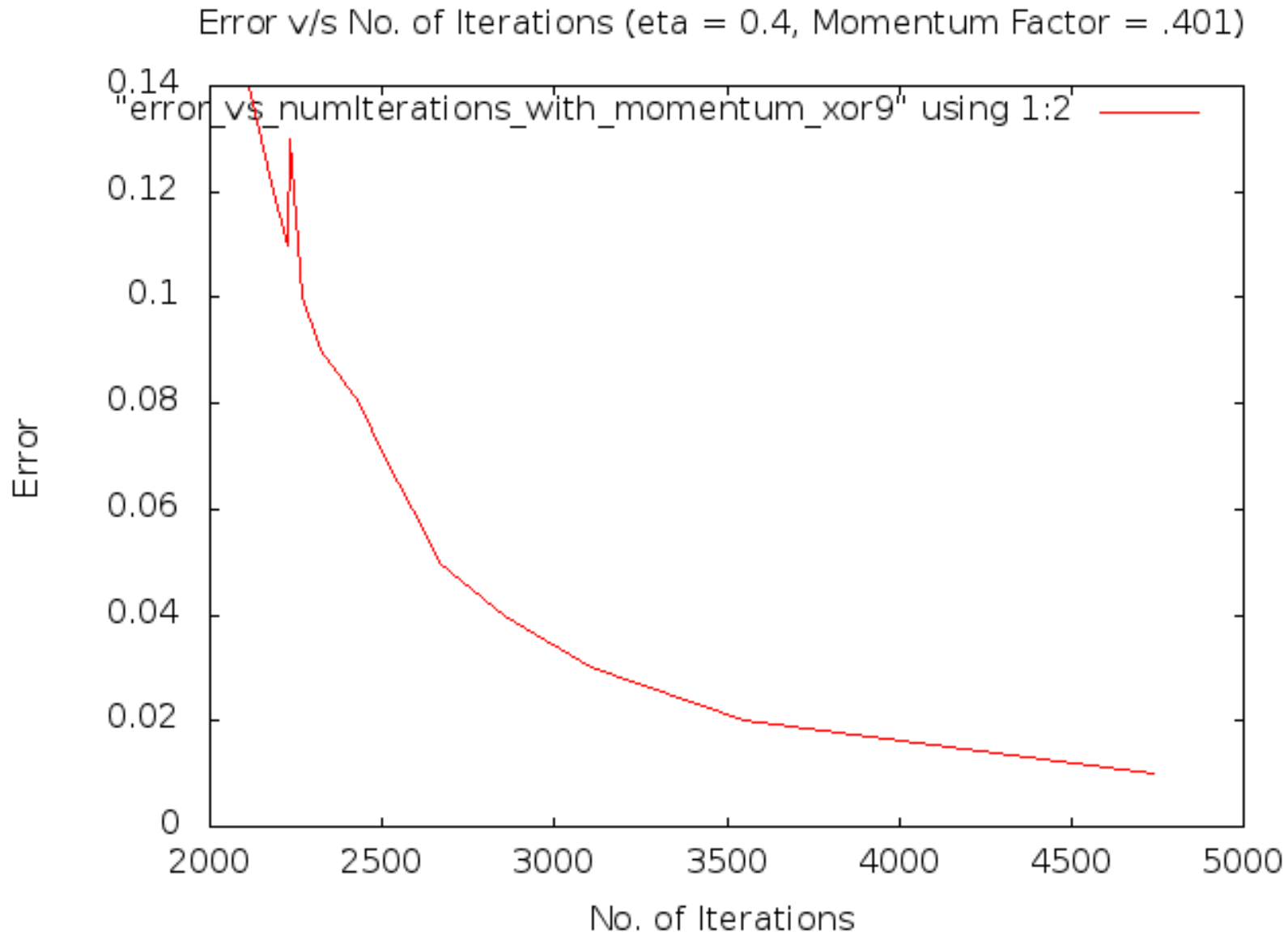
# Effect of Momentum Factor (XOR)



# Effect of Momentum Factor (XOR)

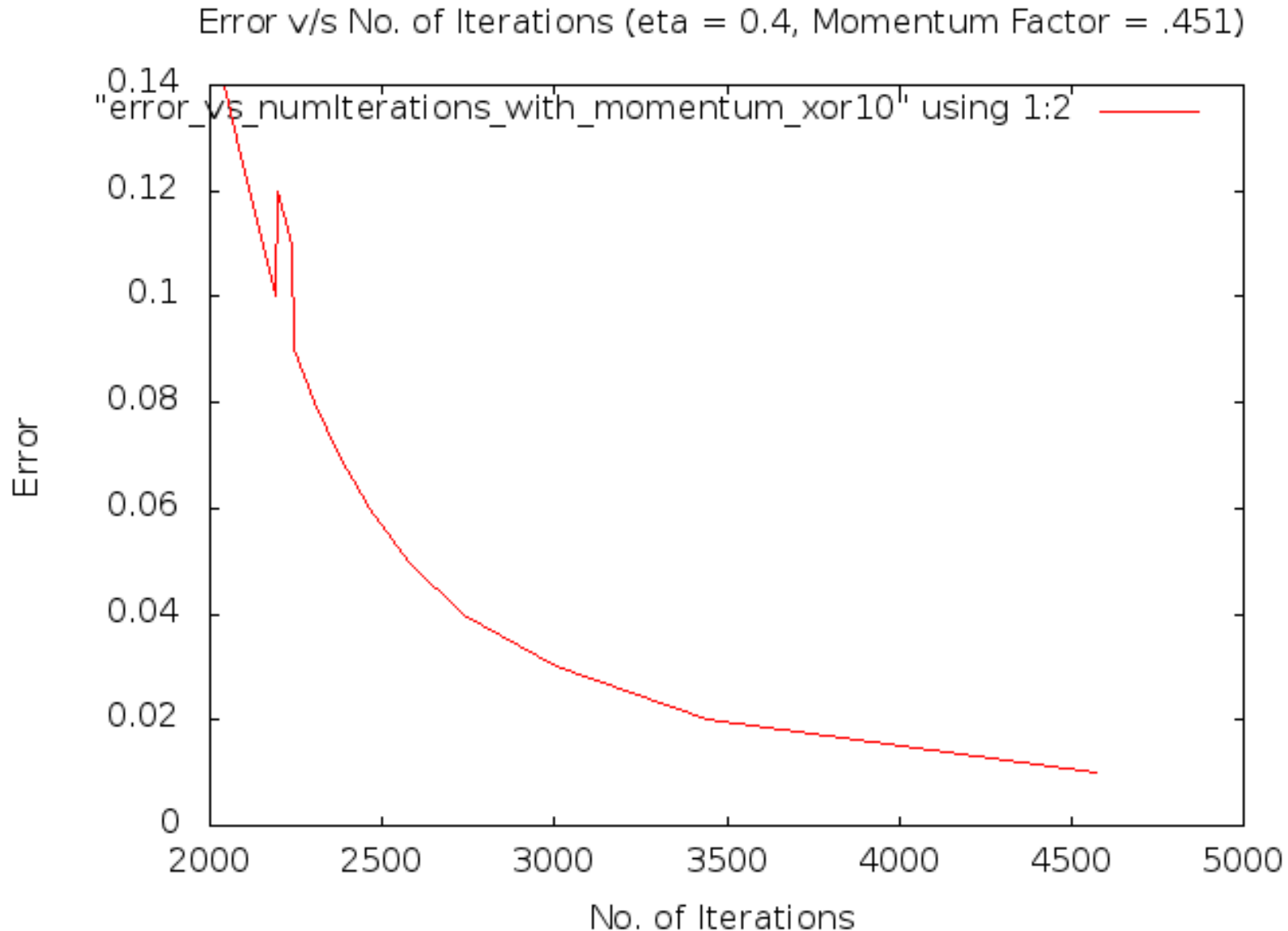


# Effect of Momentum Factor (XOR)

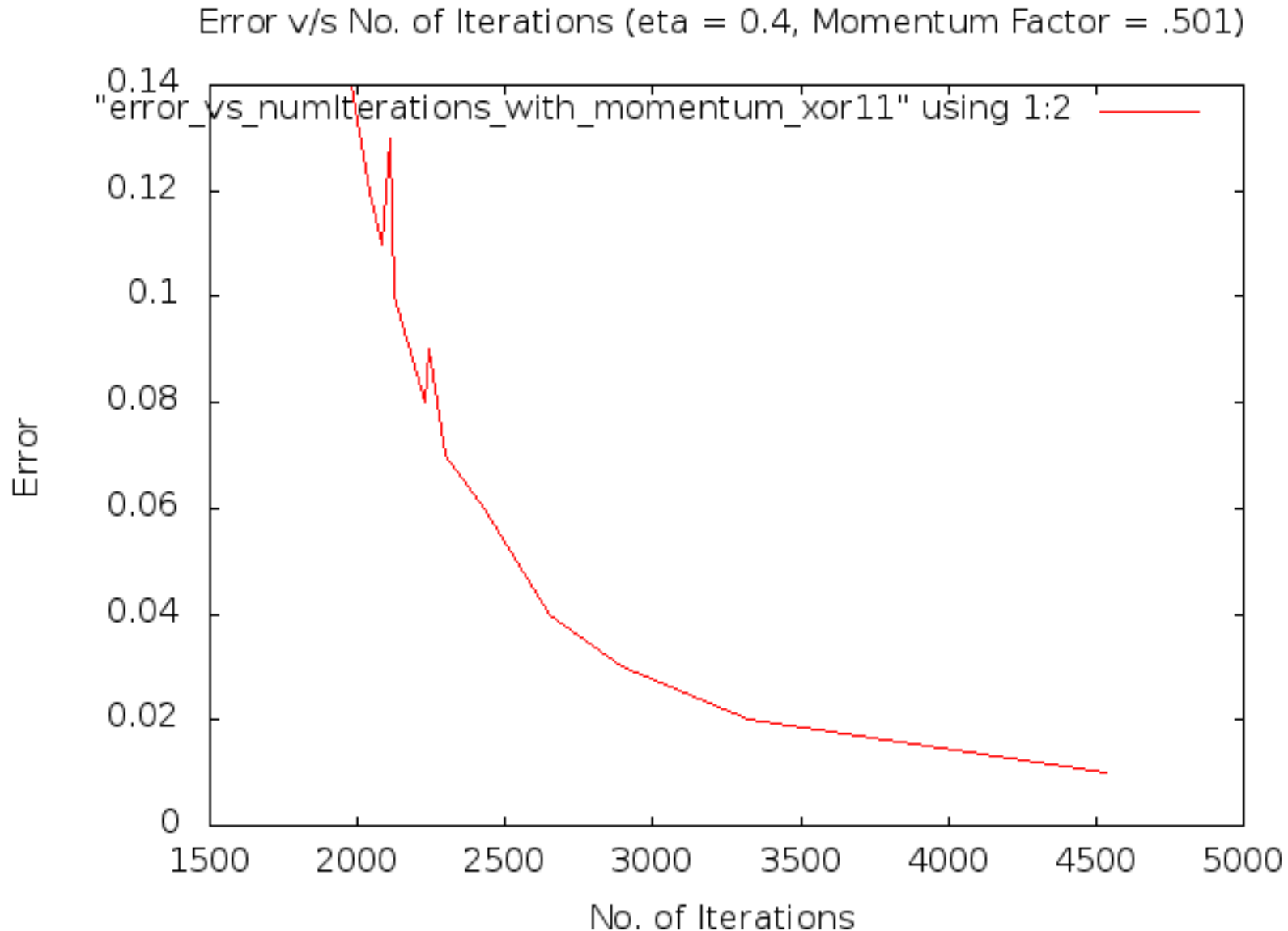




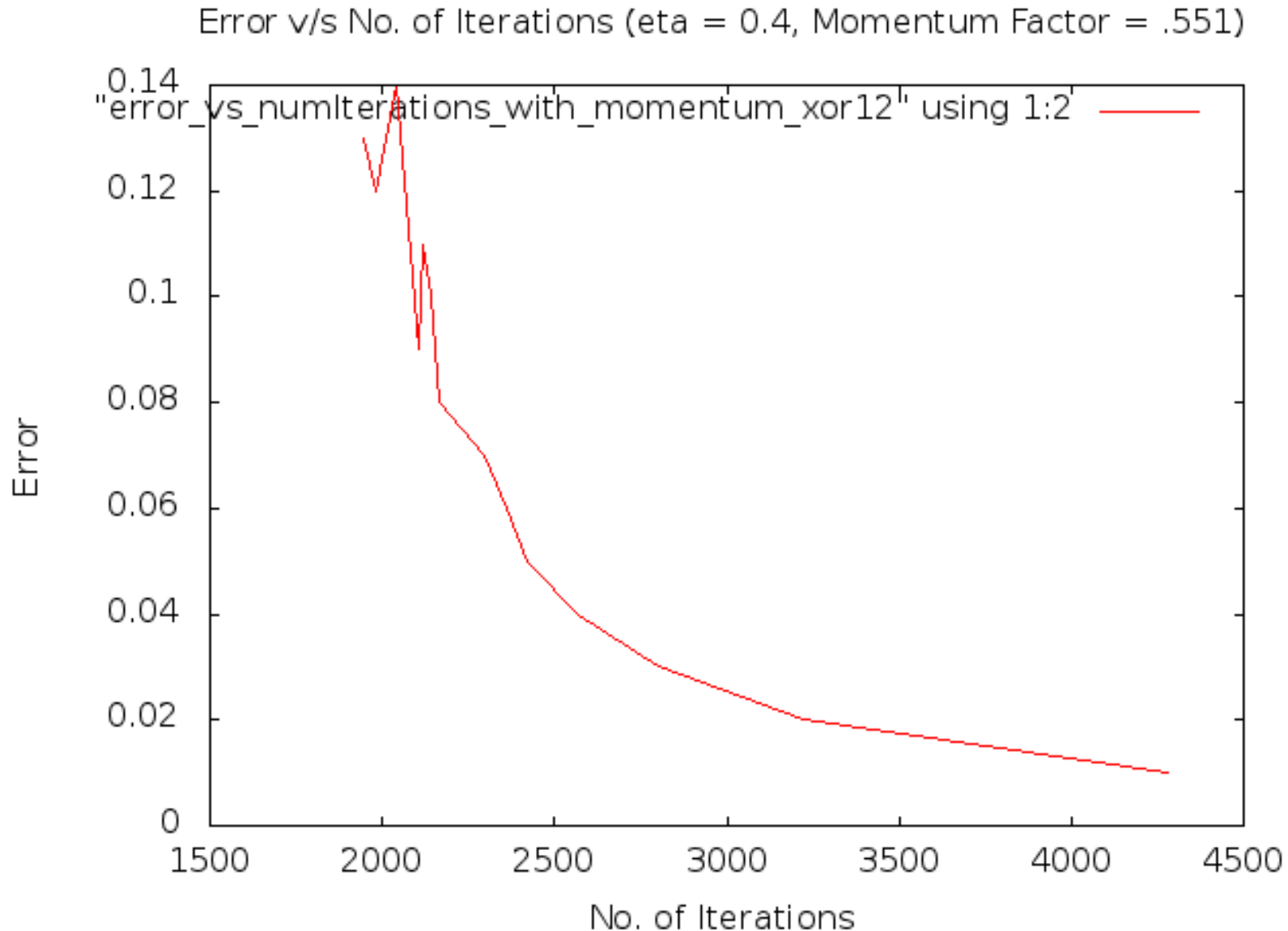
# Effect of Momentum Factor (XOR)



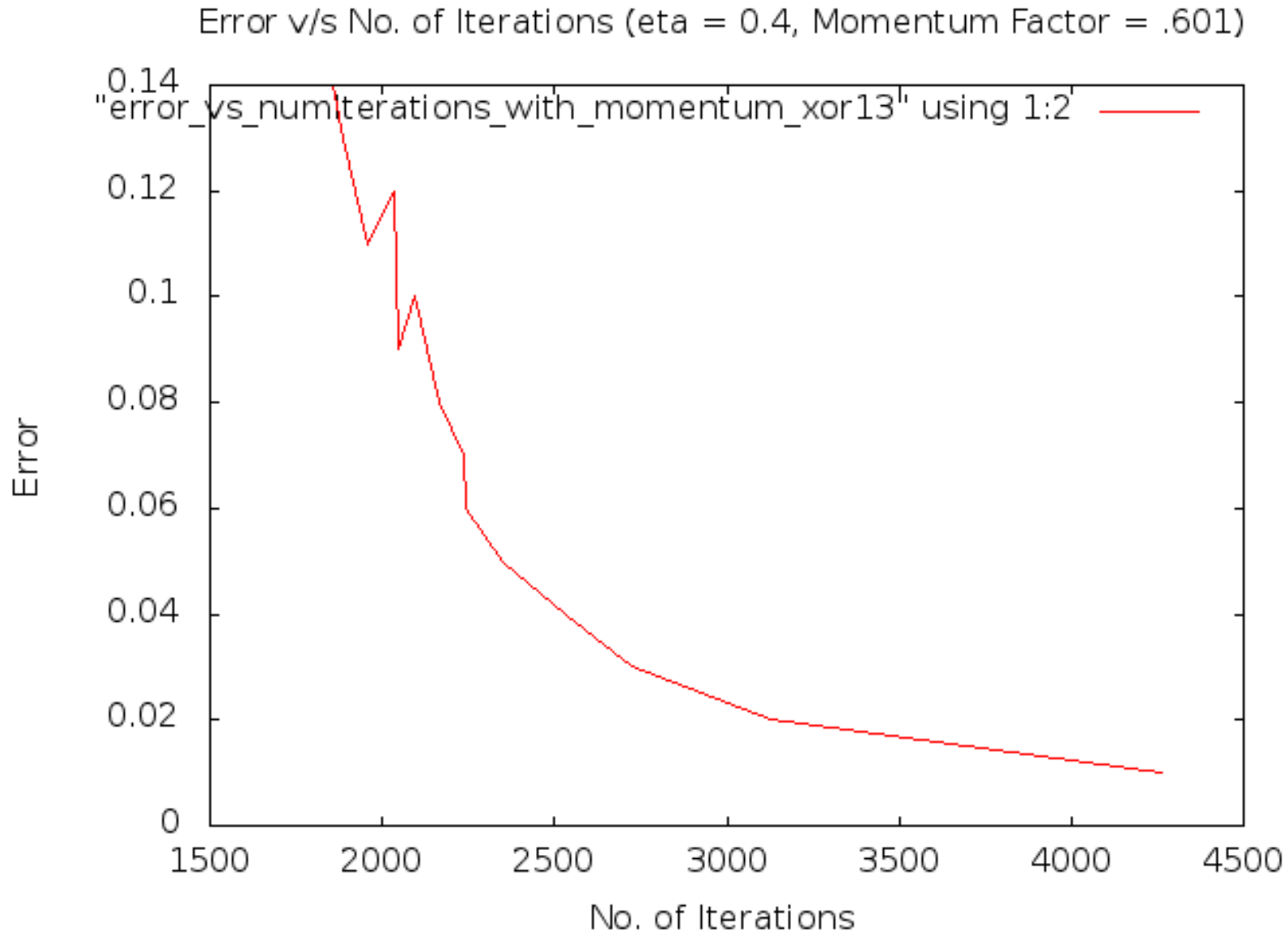
# Effect of Momentum Factor (XOR)



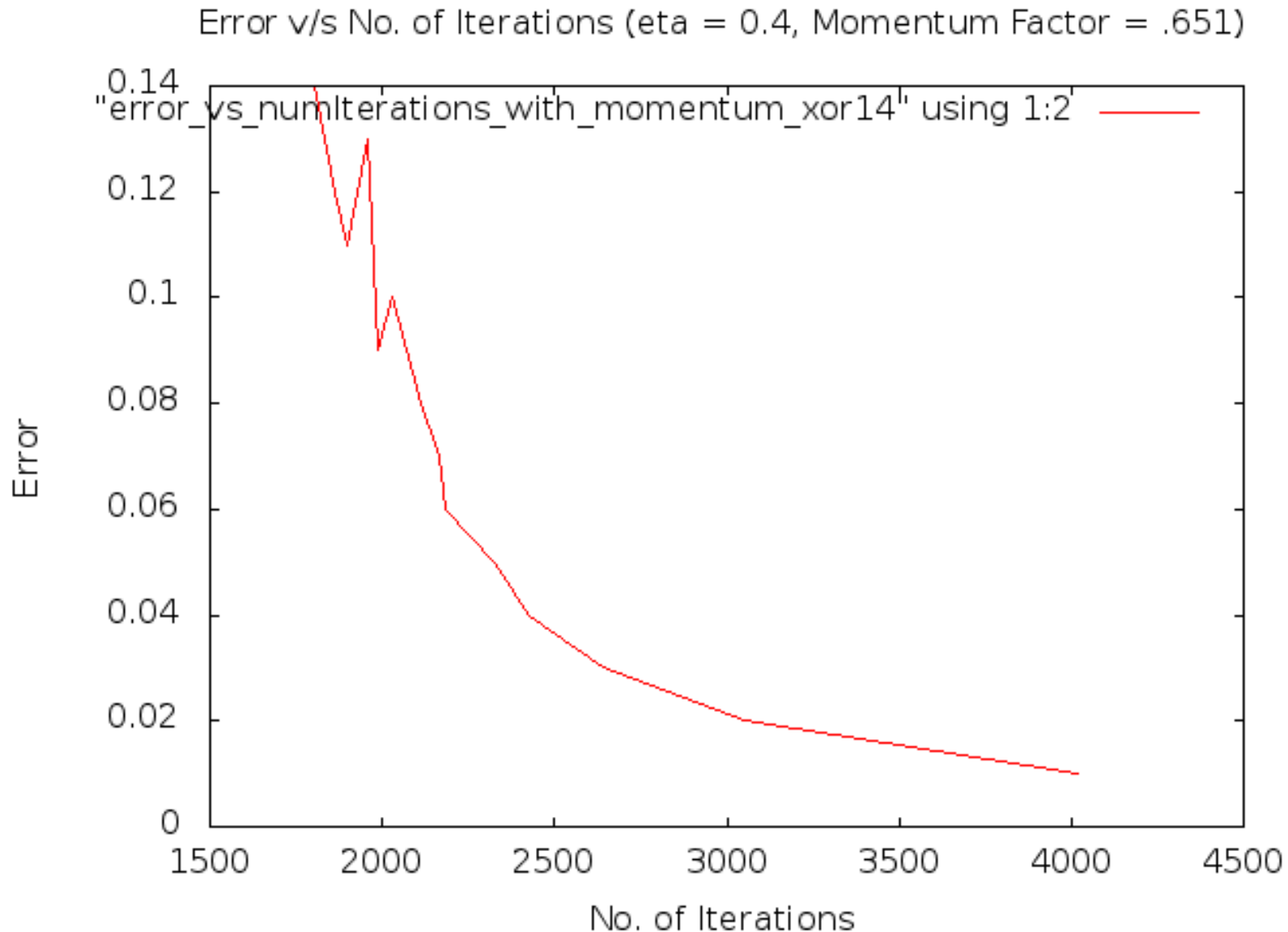
# Effect of Momentum Factor (XOR)



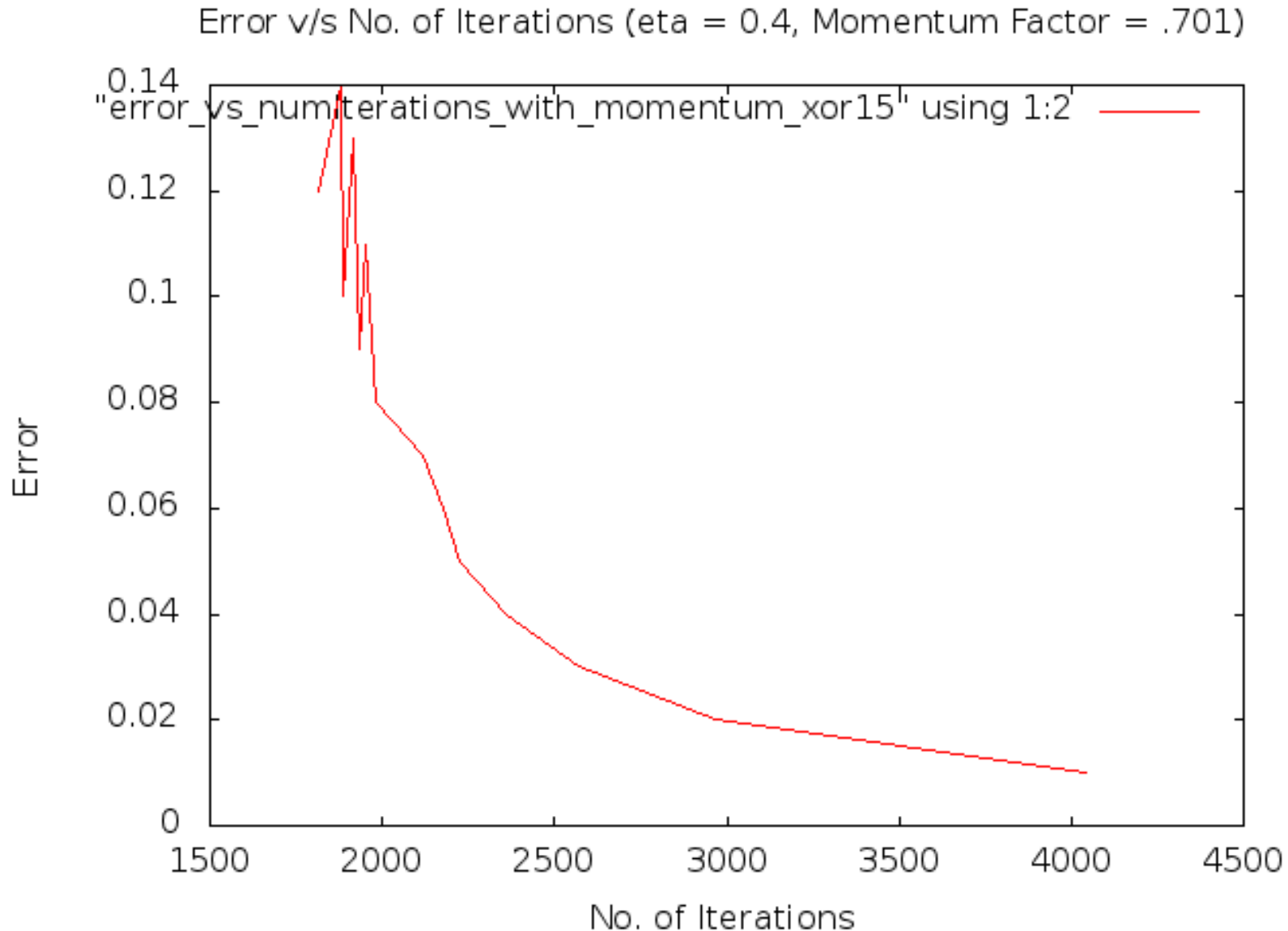
# Effect of Momentum Factor (XOR)



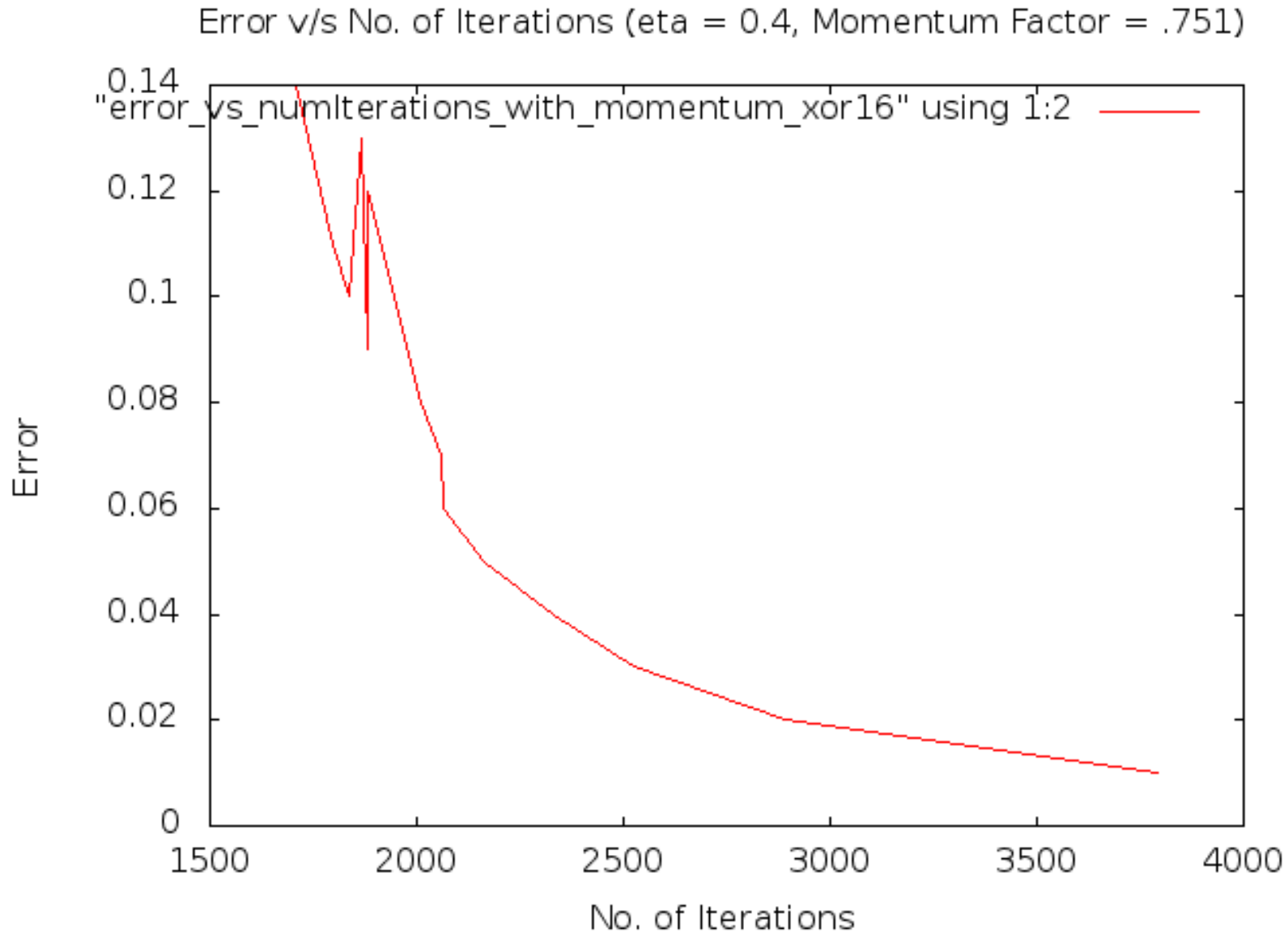
# Effect of Momentum Factor (XOR)



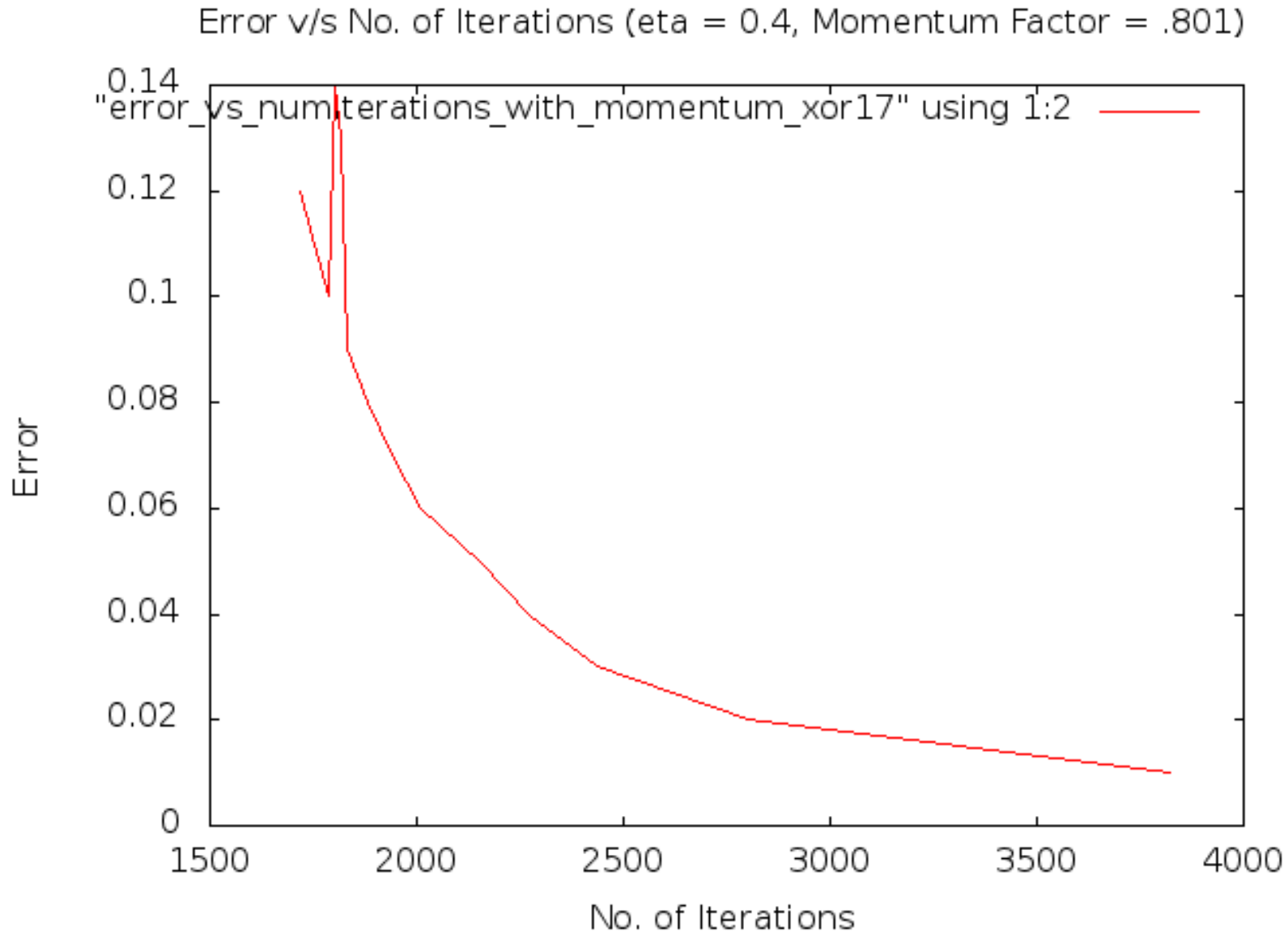
# Effect of Momentum Factor (XOR)



# Effect of Momentum Factor (XOR)

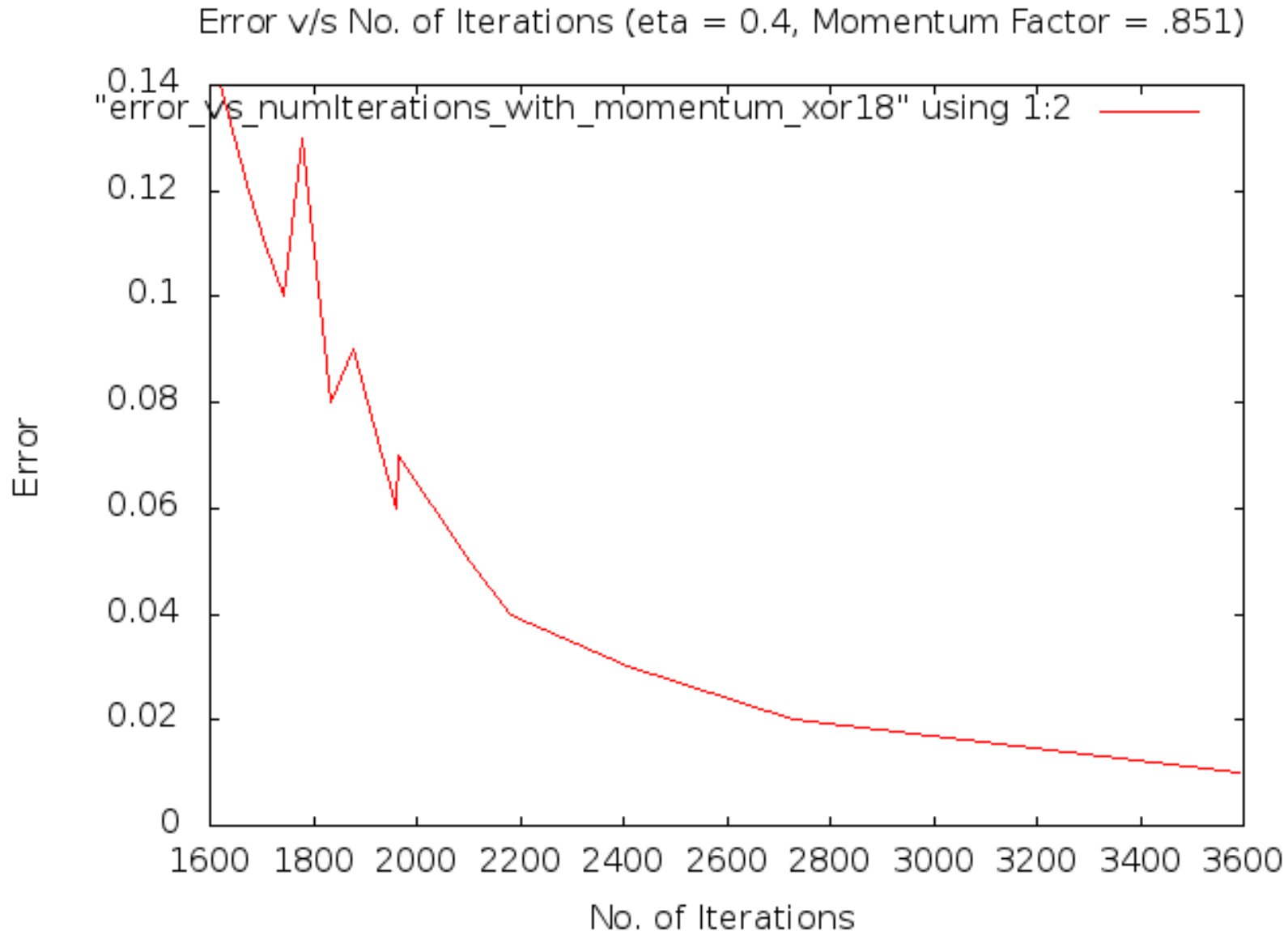


# Effect of Momentum Factor (XOR)

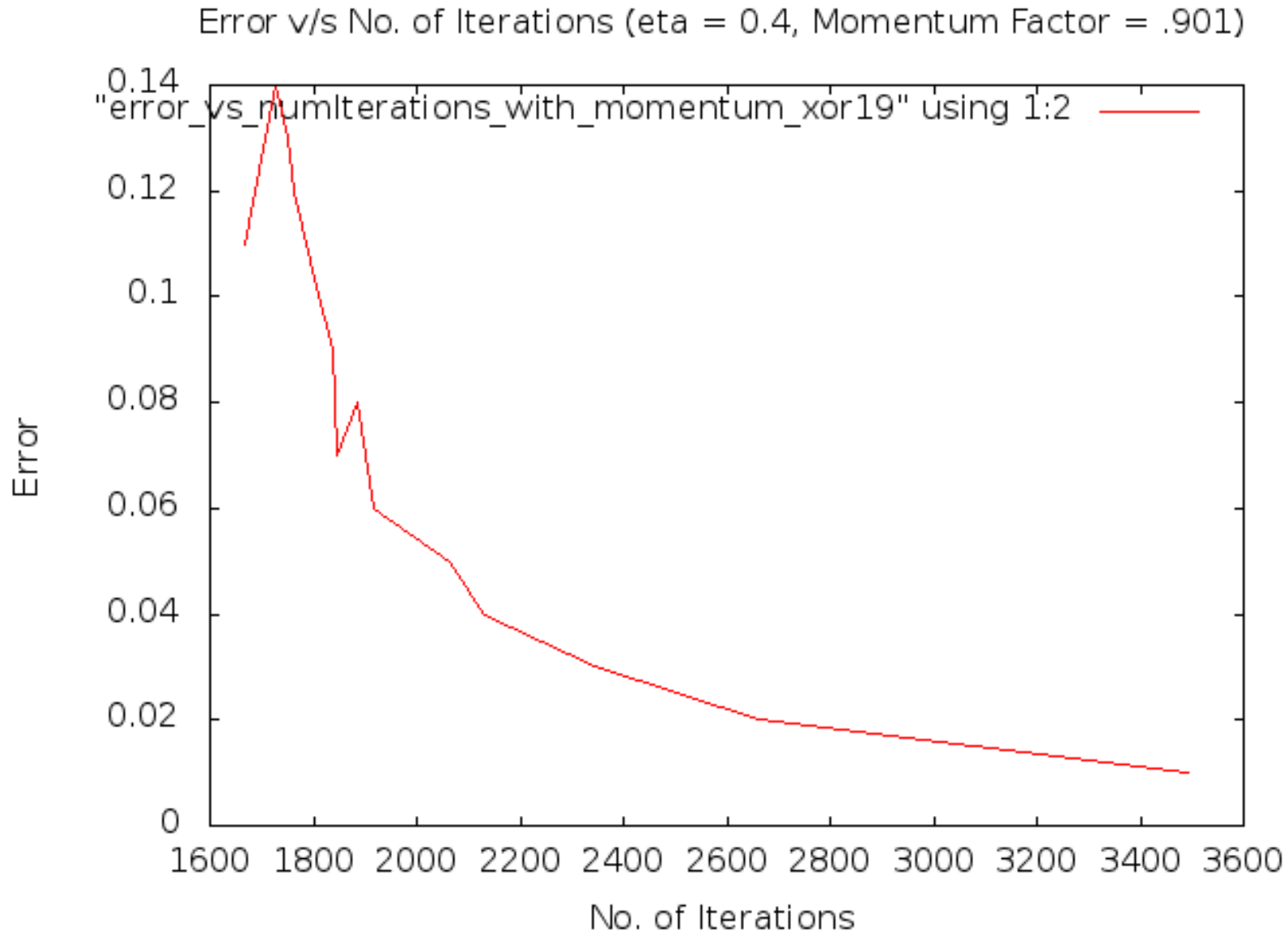




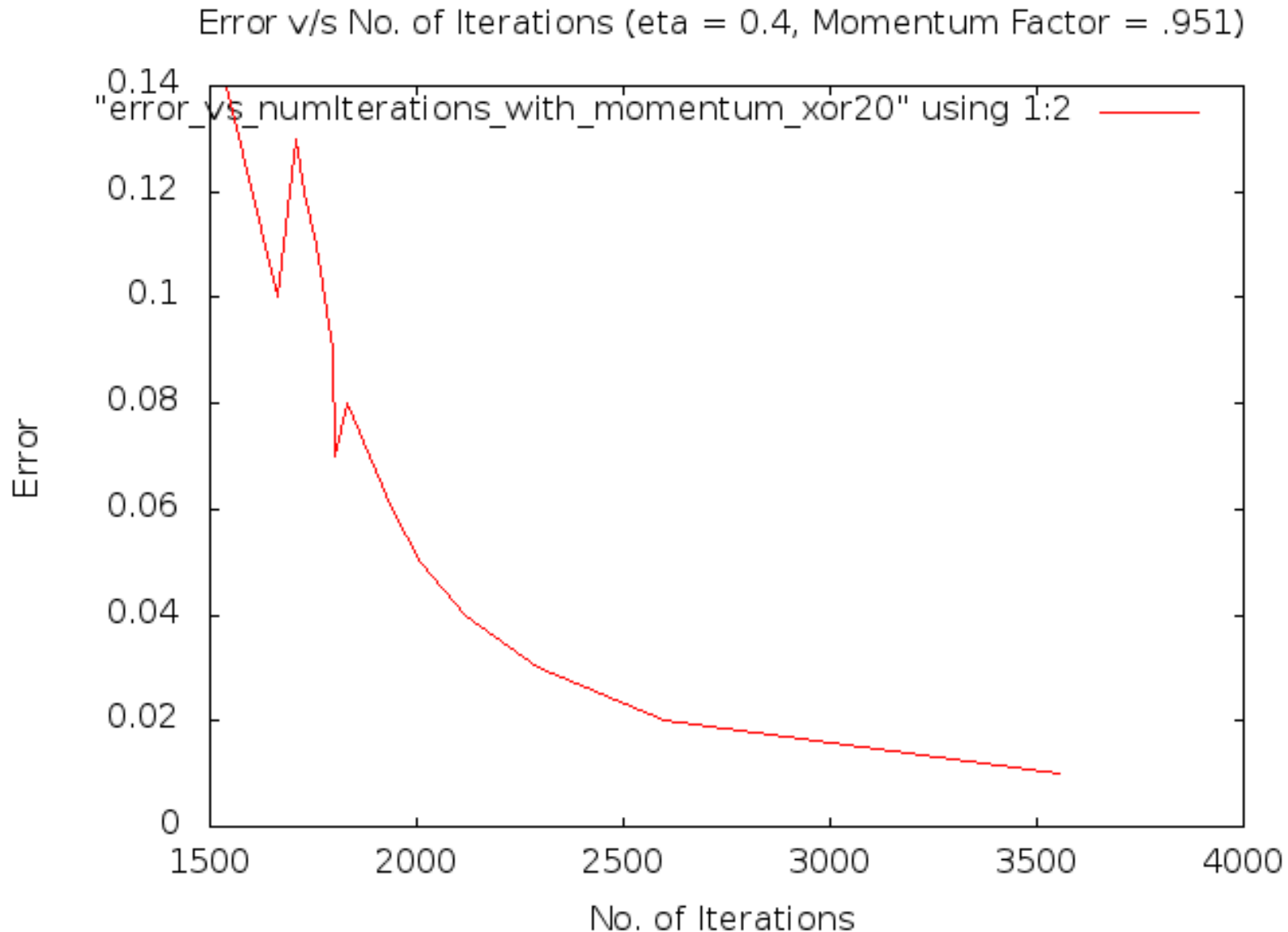
# Effect of Momentum Factor (XOR)



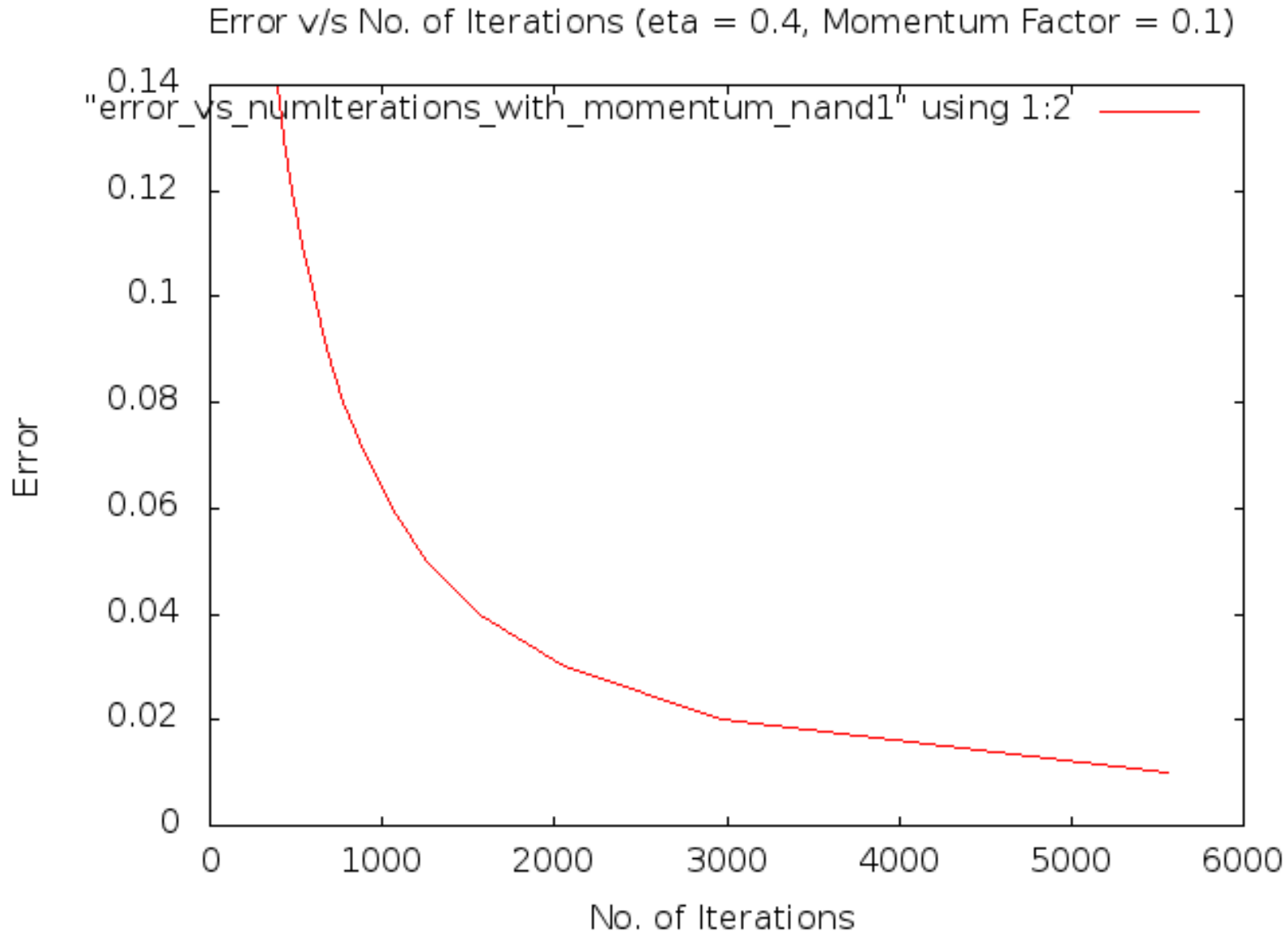
# Effect of Momentum Factor (XOR)



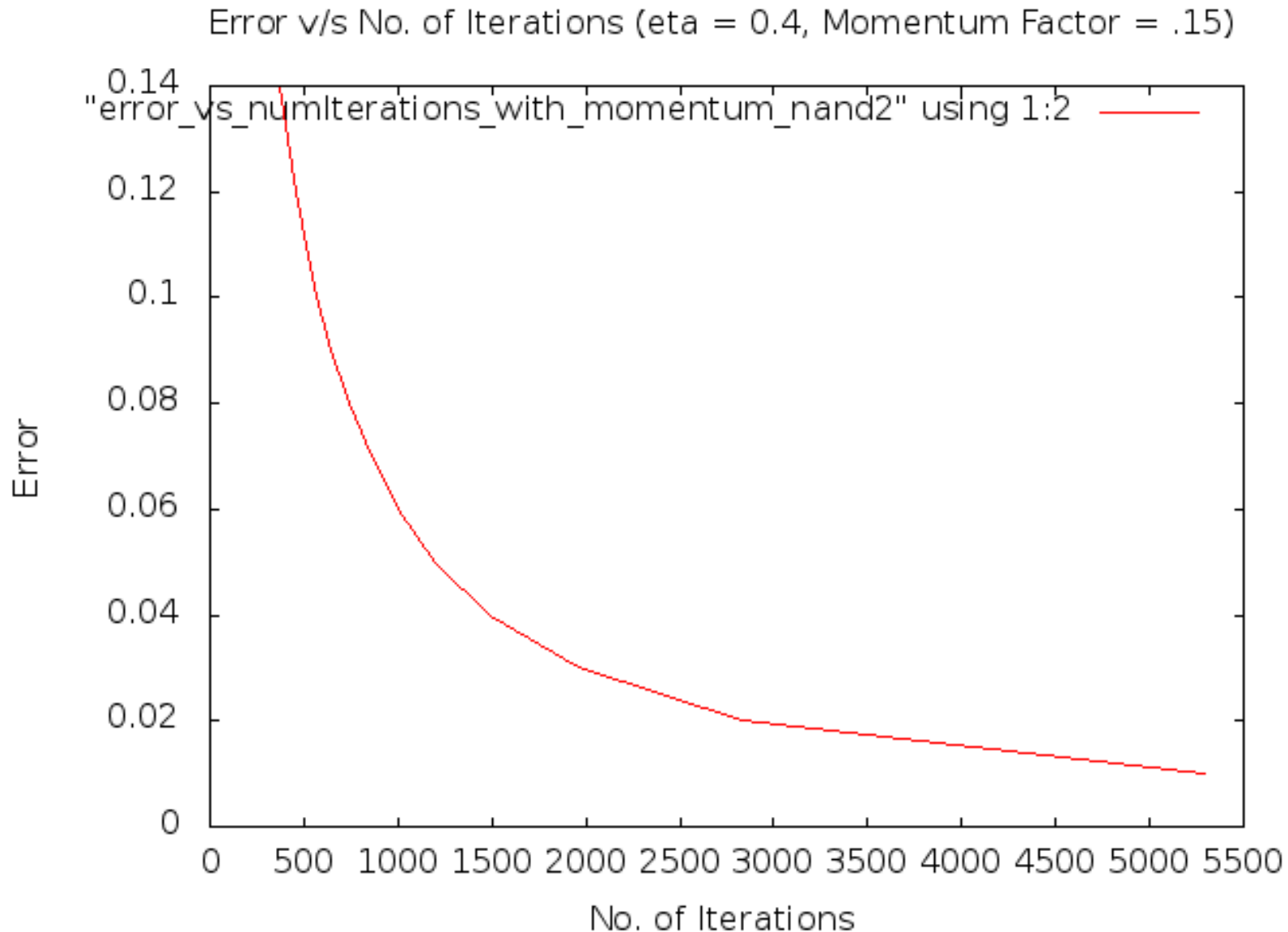
# Effect of Momentum Factor (XOR)



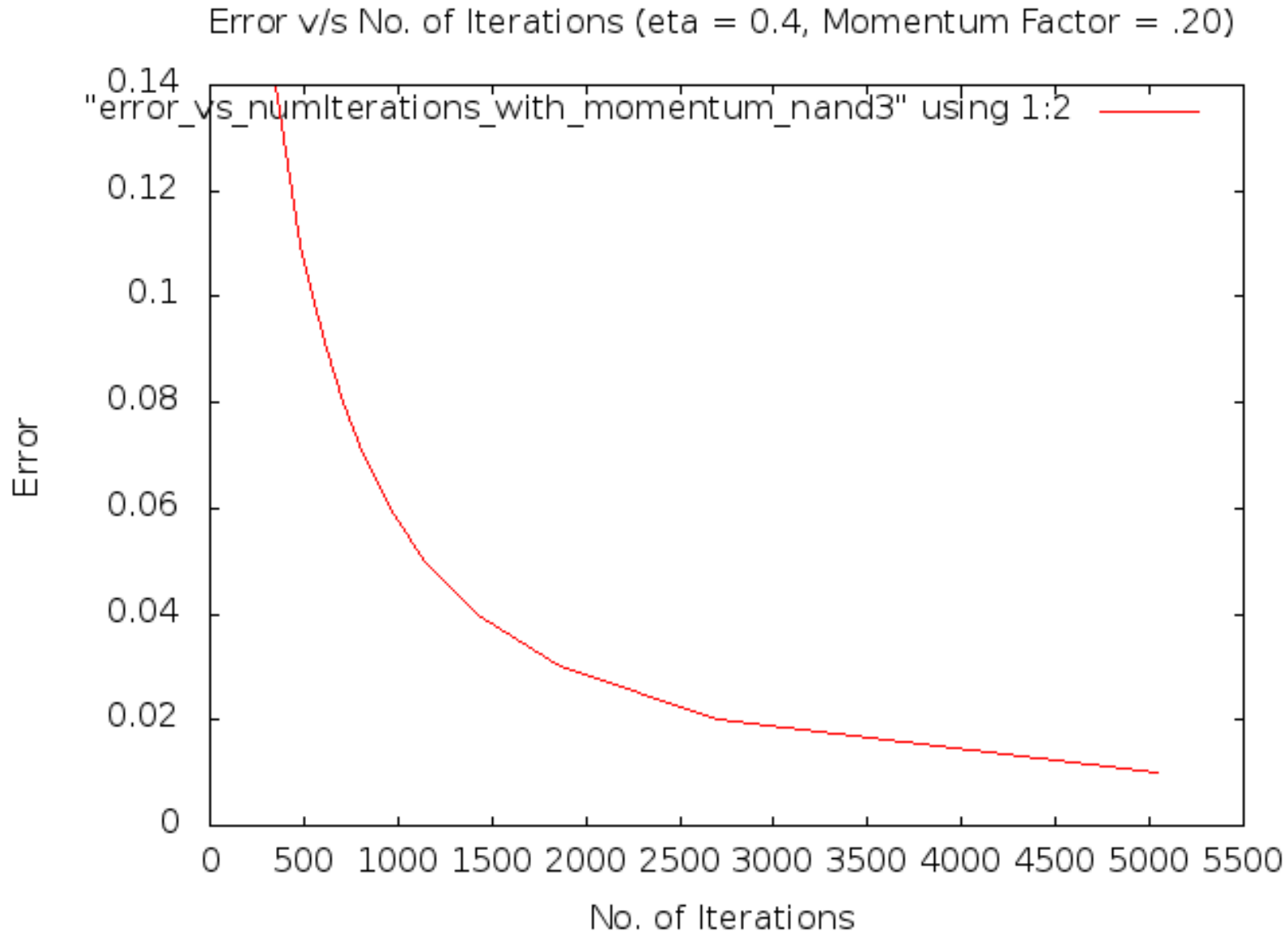
# Effect of Momentum Factor (NAND)



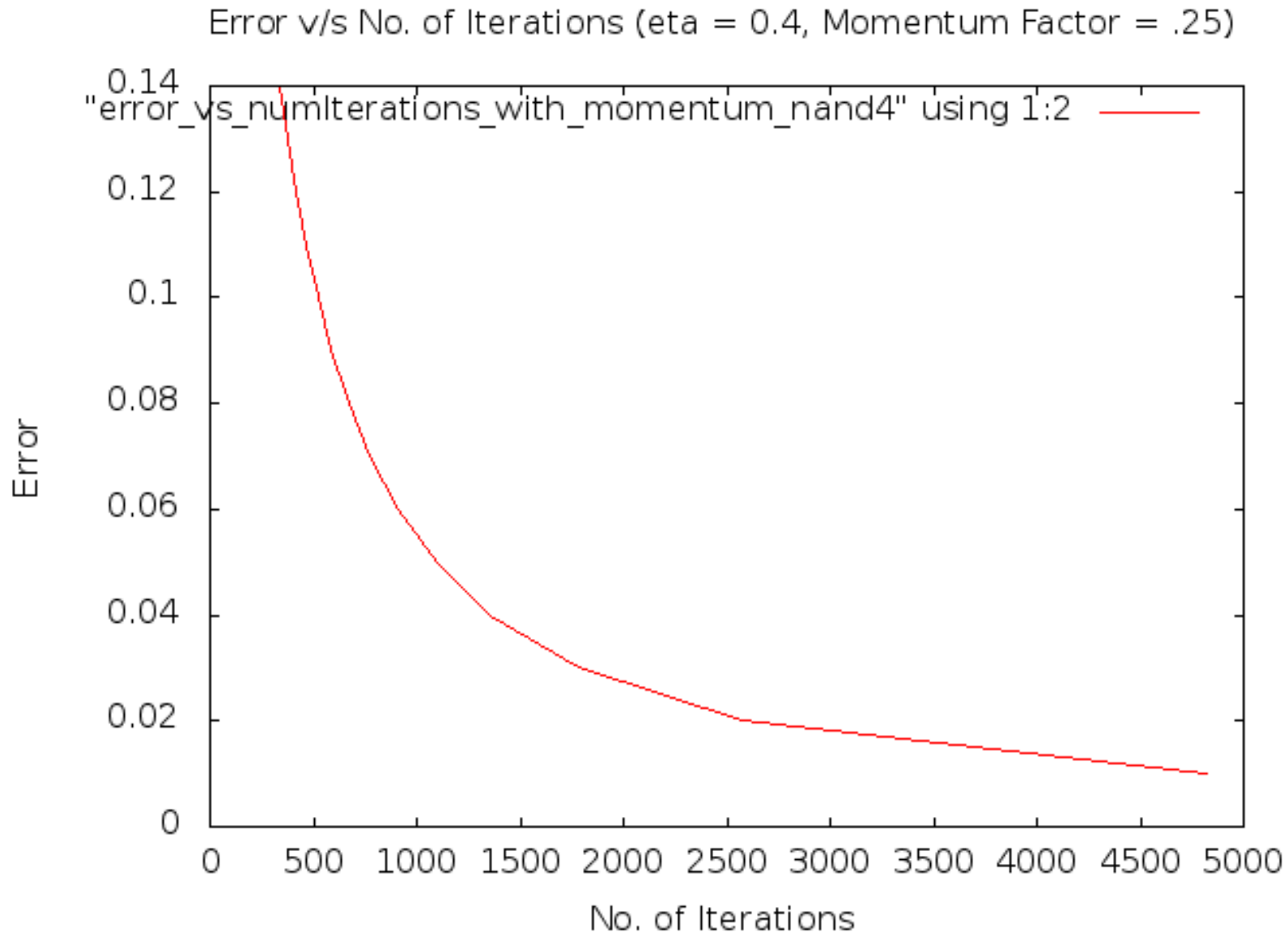
# Effect of Momentum Factor (NAND)



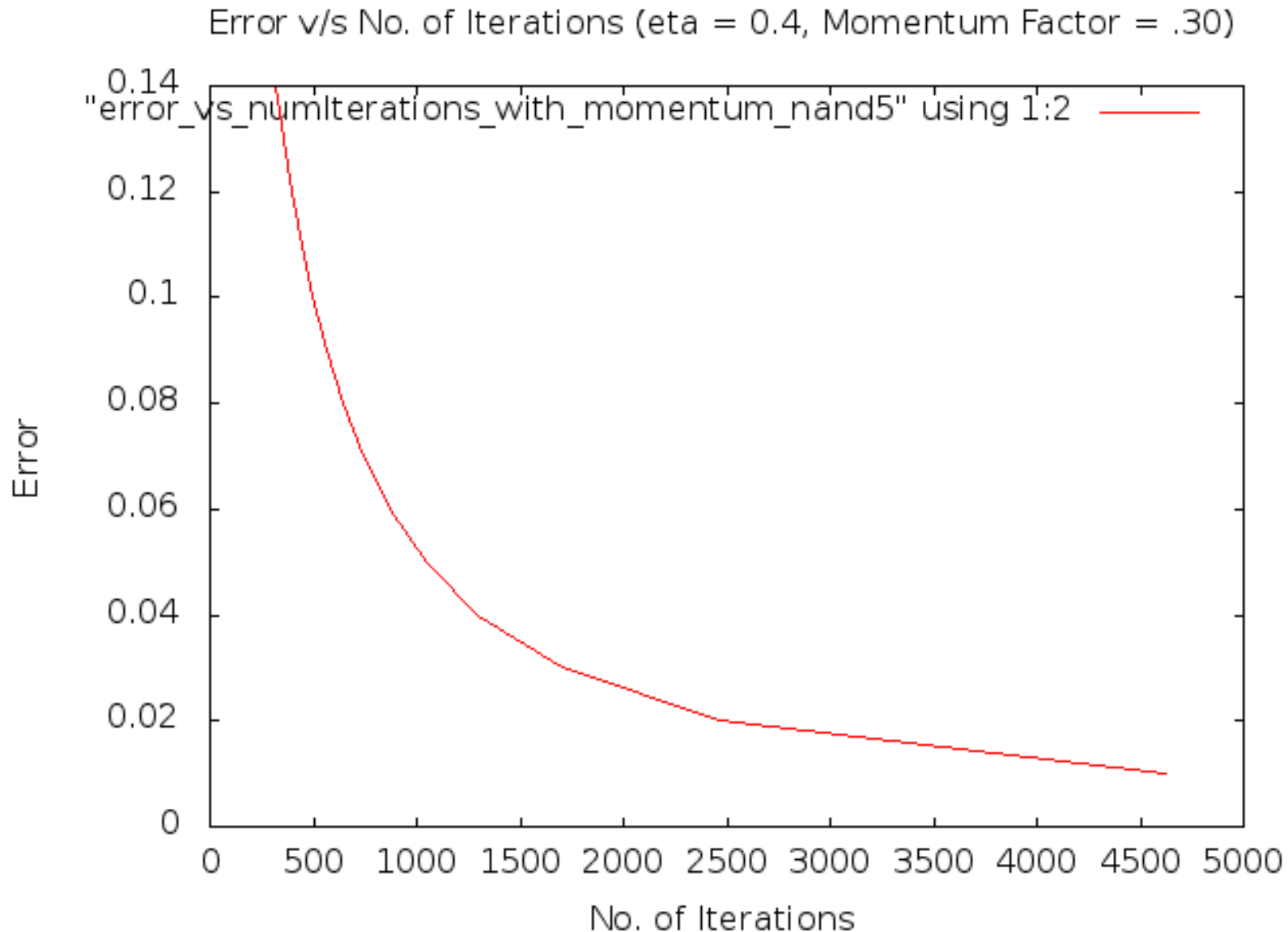
# Effect of Momentum Factor (NAND)



# Effect of Momentum Factor (NAND)

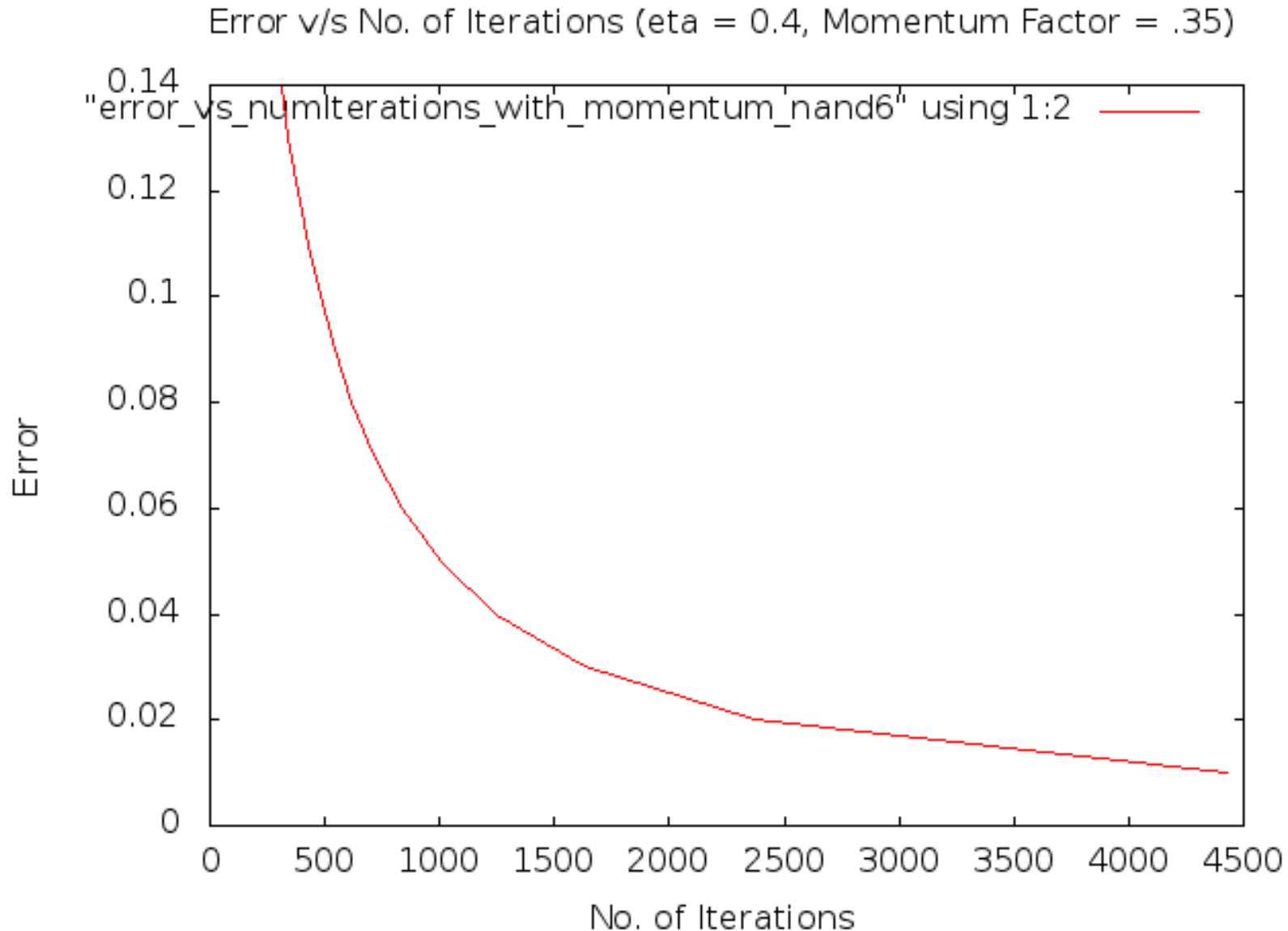


# Effect of Momentum Factor (NAND)

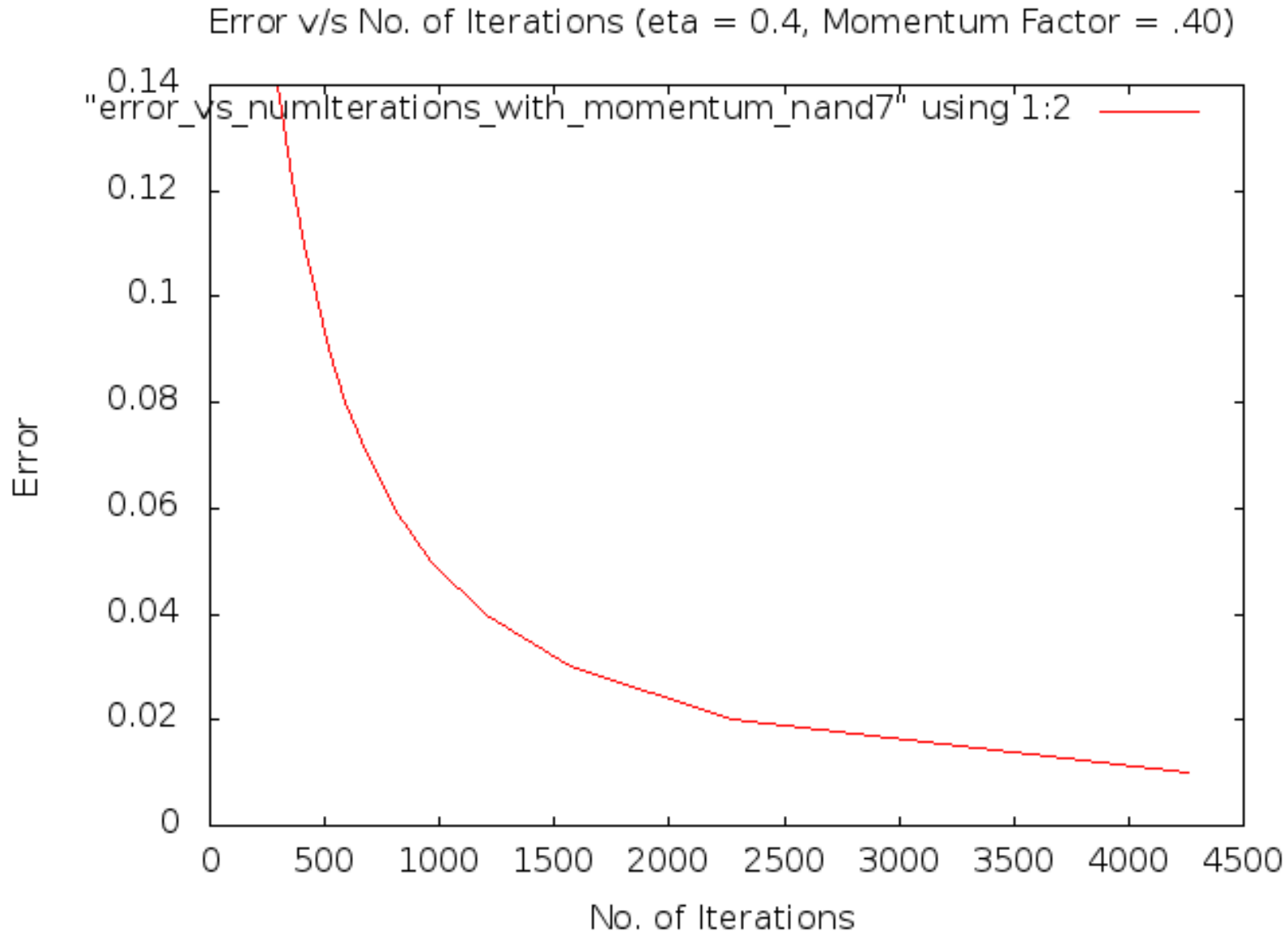




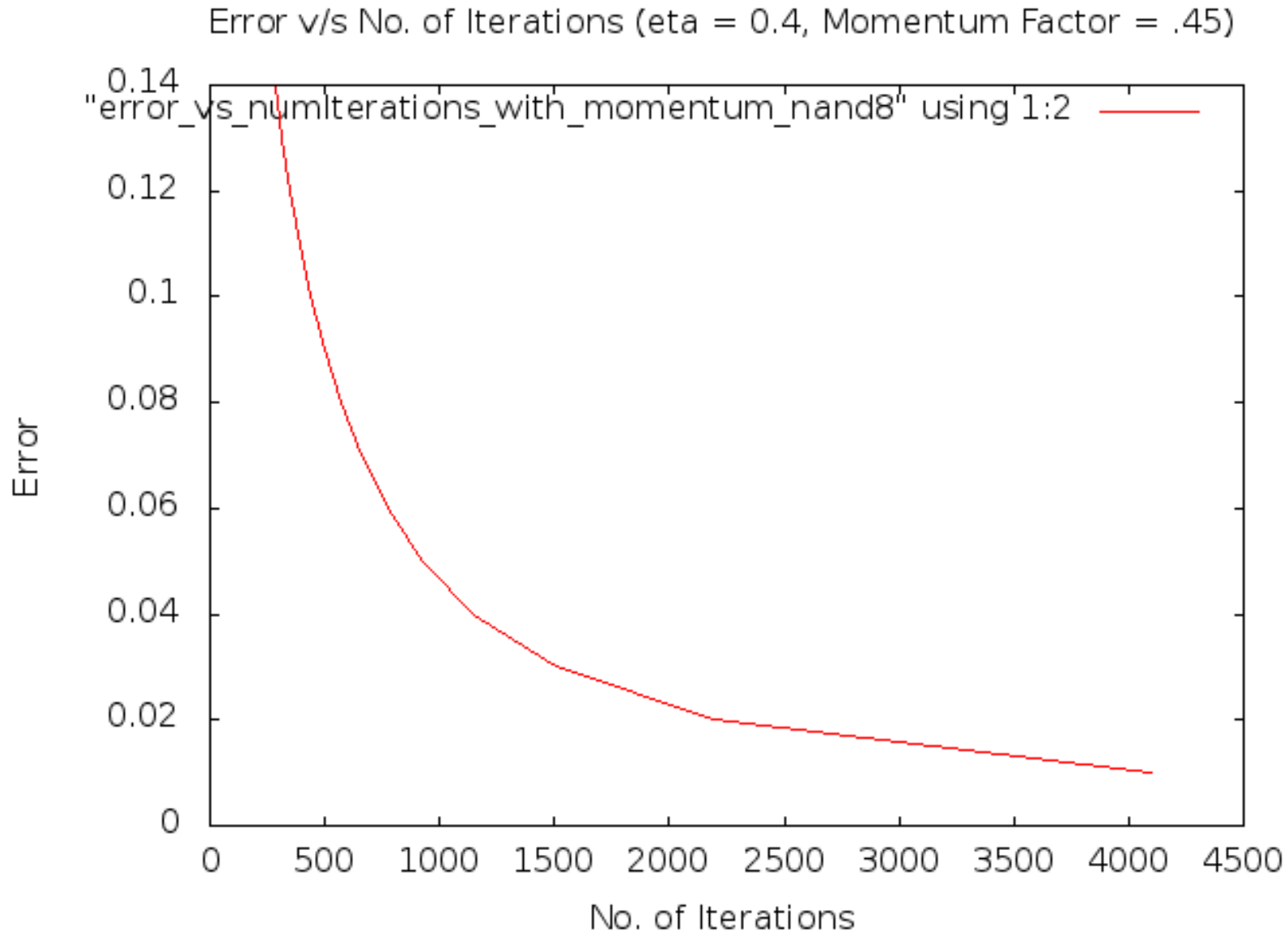
# Effect of Momentum Factor (NAND)



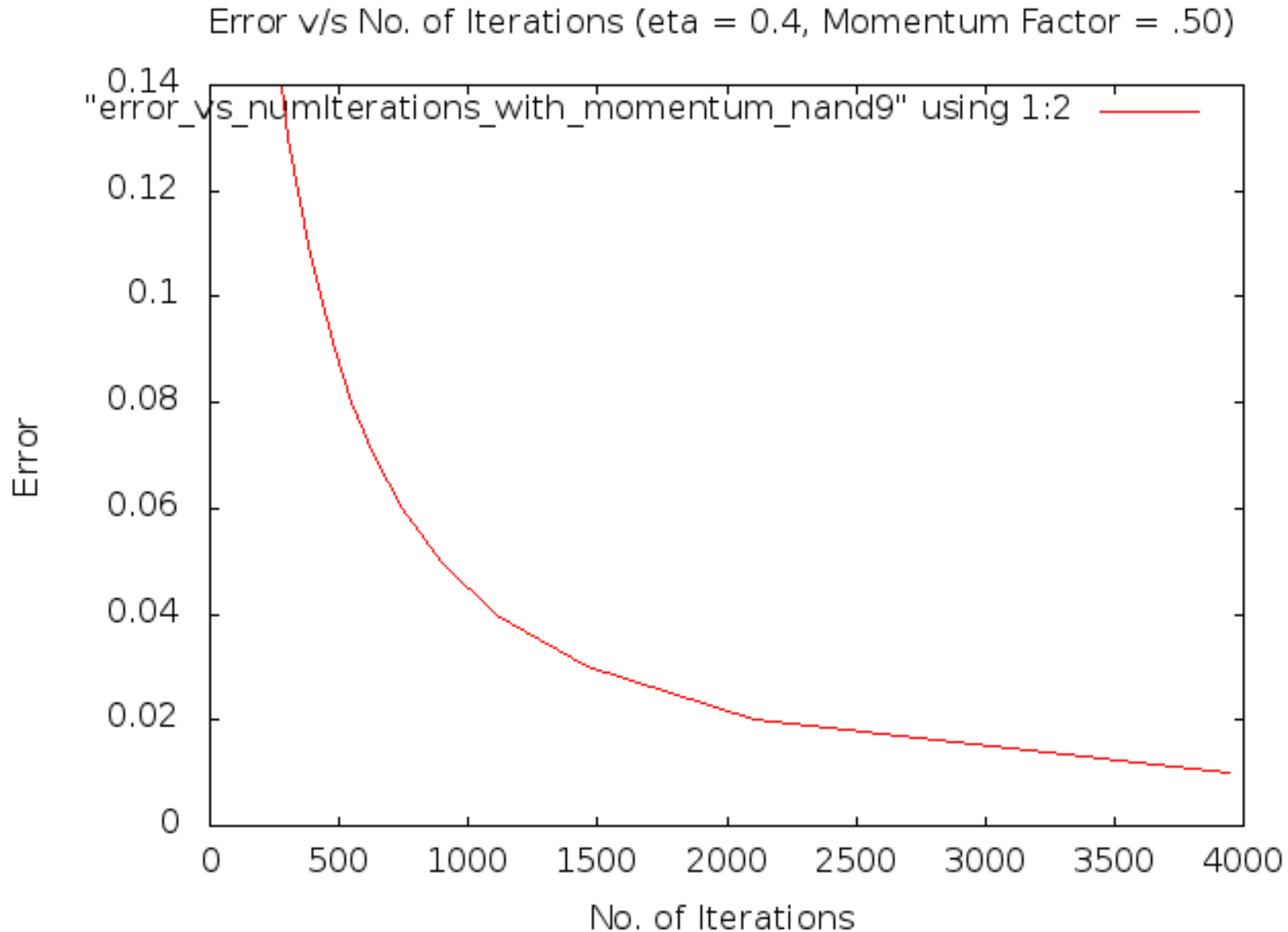
# Effect of Momentum Factor (NAND)



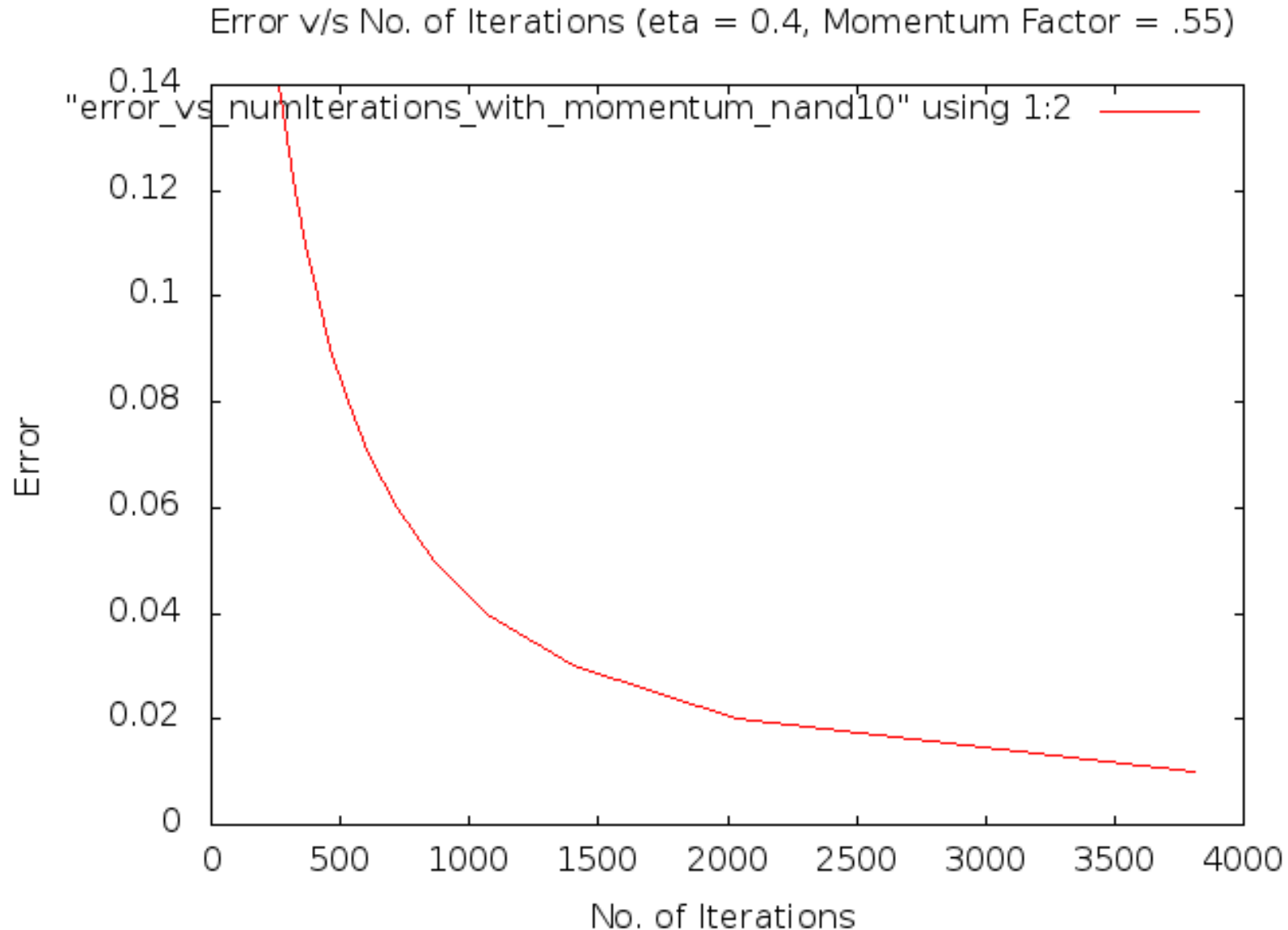
# Effect of Momentum Factor (NAND)



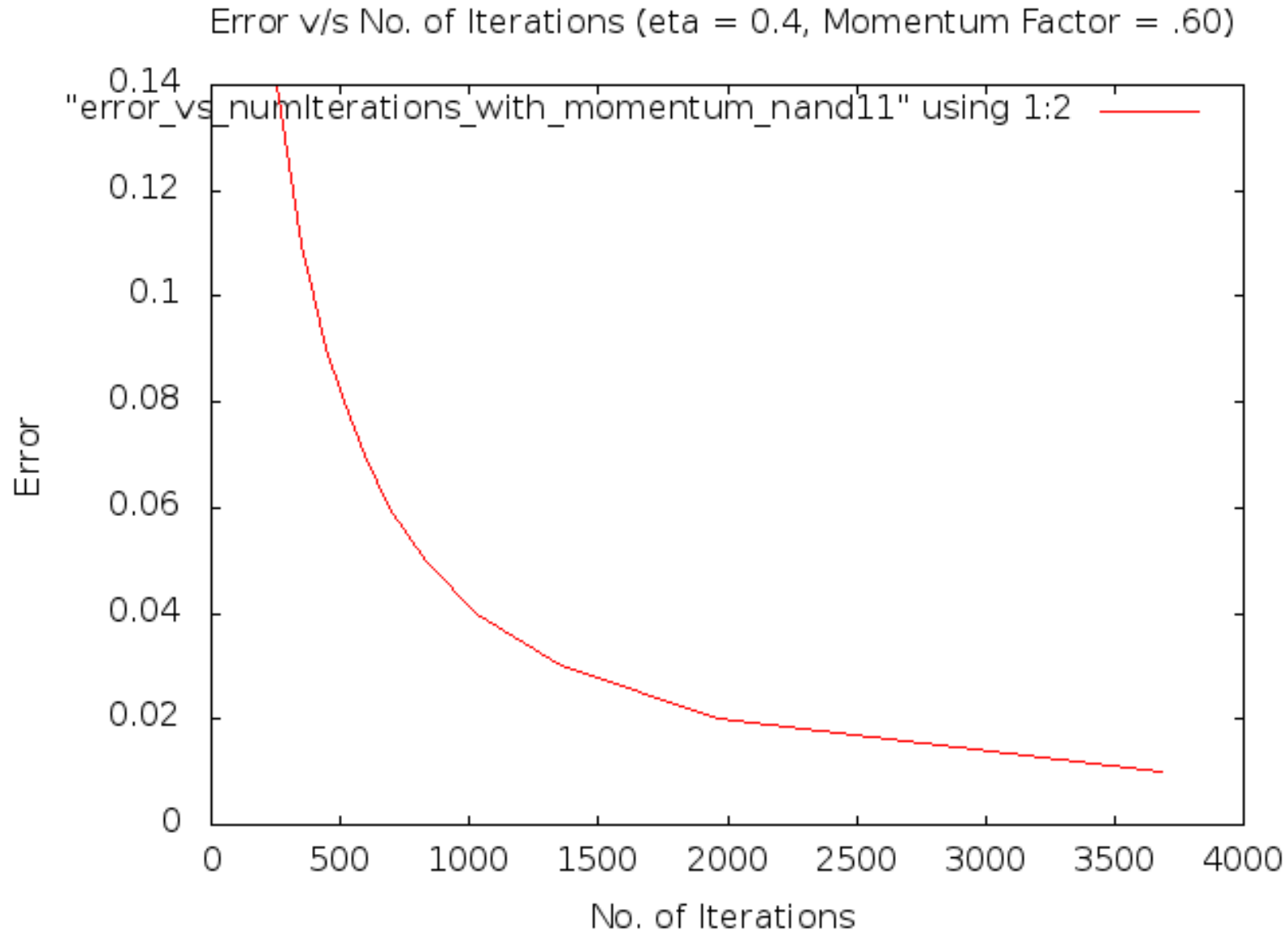
# Effect of Momentum Factor (NAND)



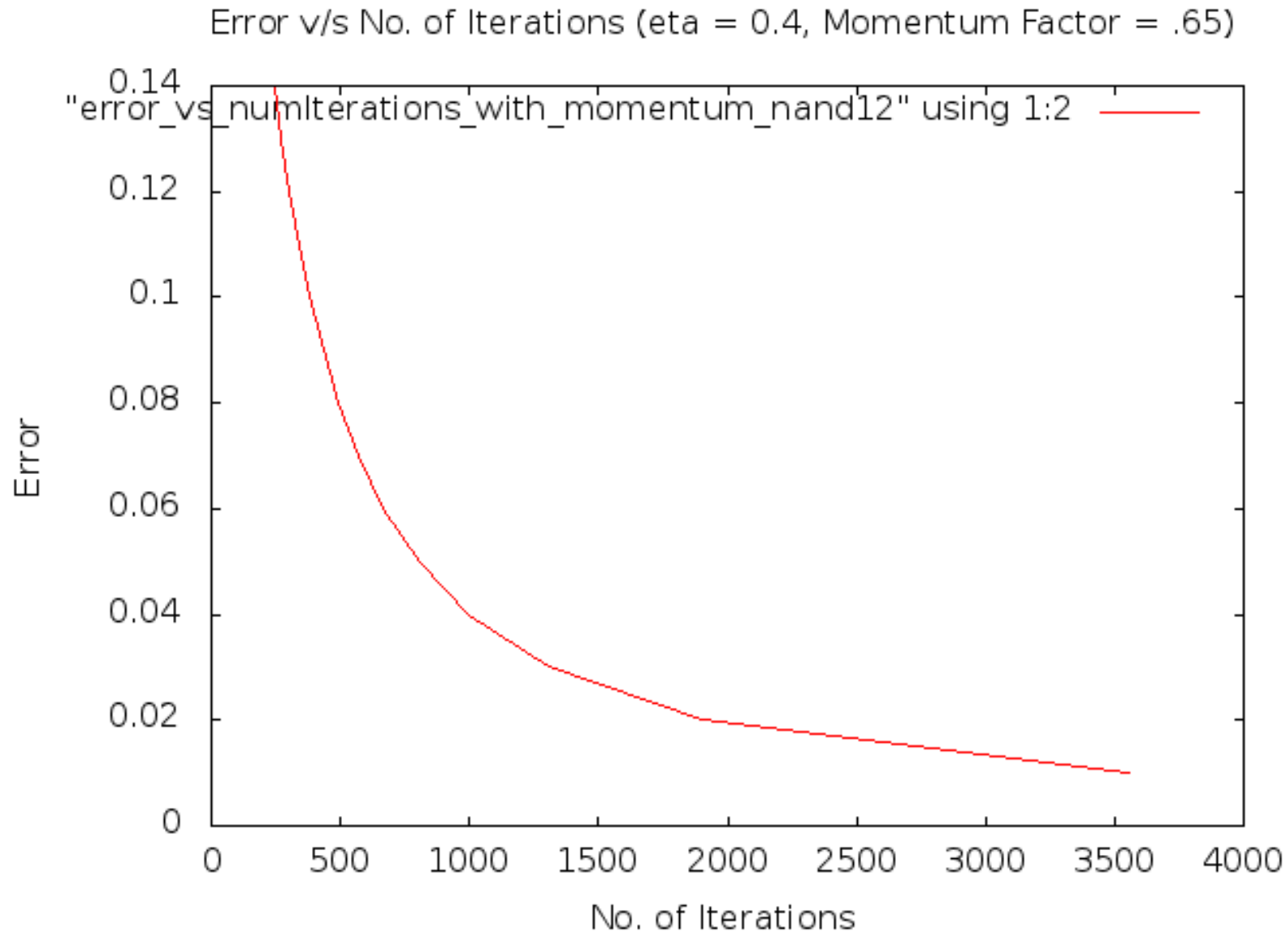
# Effect of Momentum Factor (NAND)



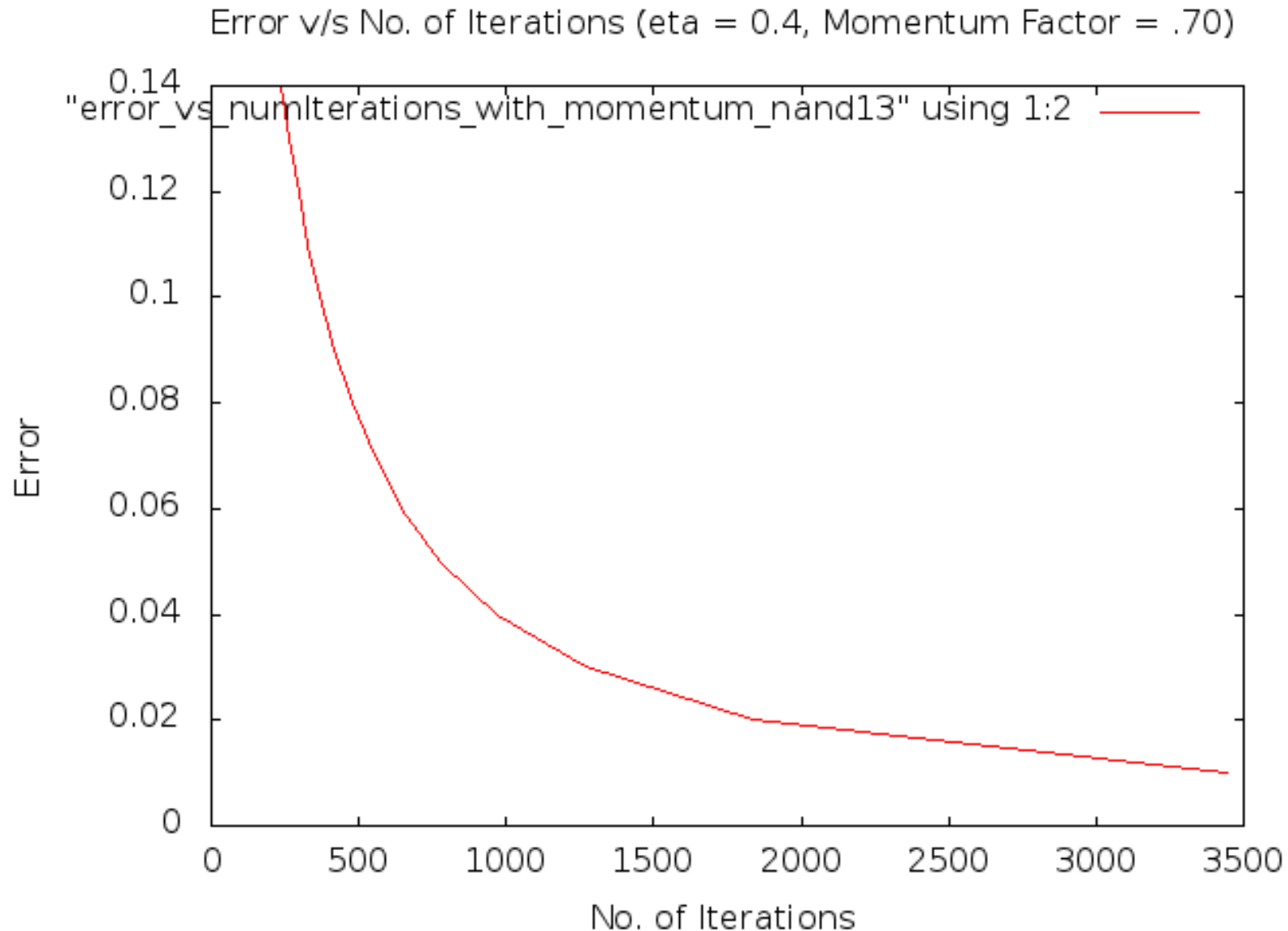
# Effect of Momentum Factor (NAND)



# Effect of Momentum Factor (NAND)

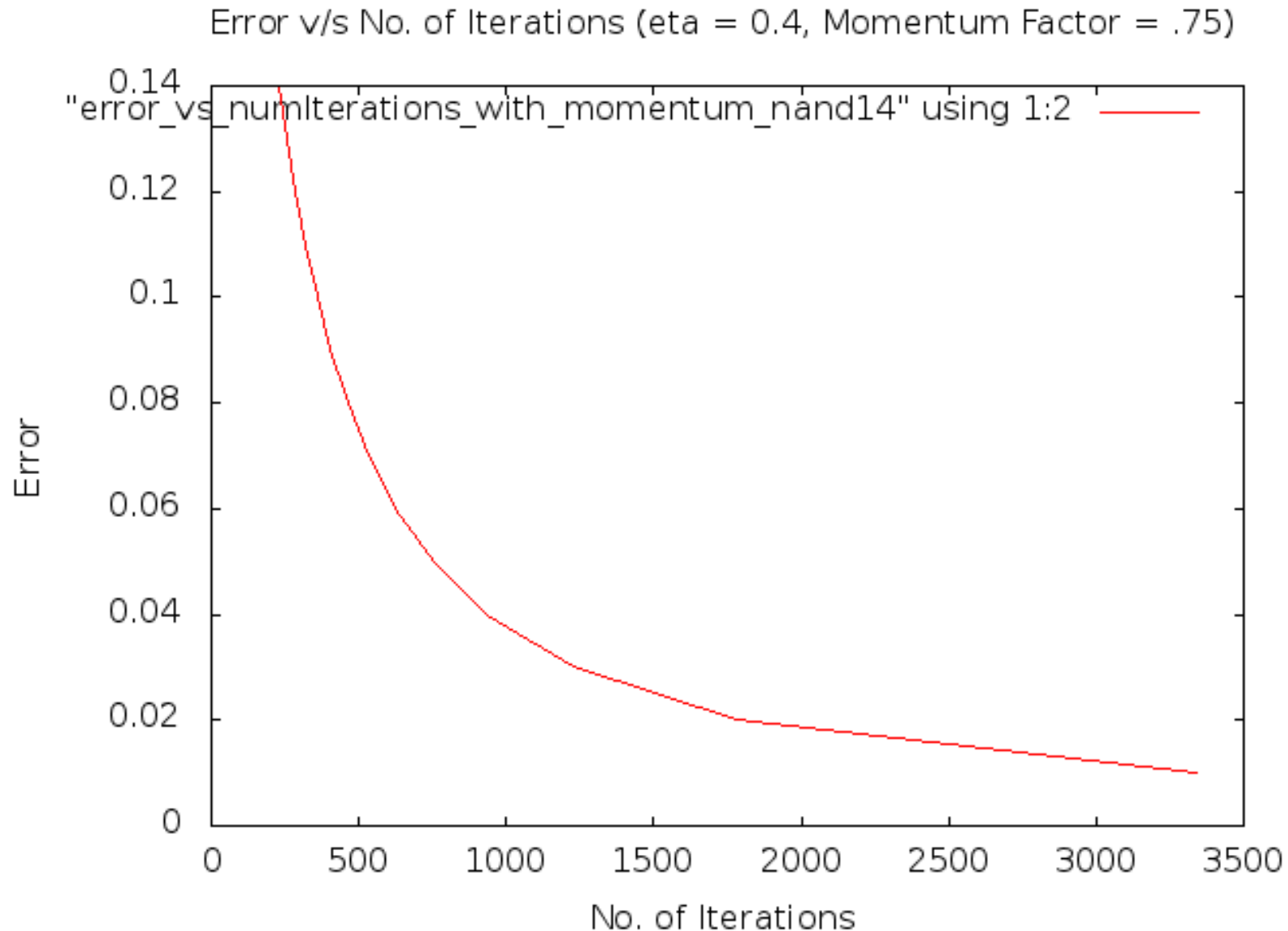


# Effect of Momentum Factor (NAND)

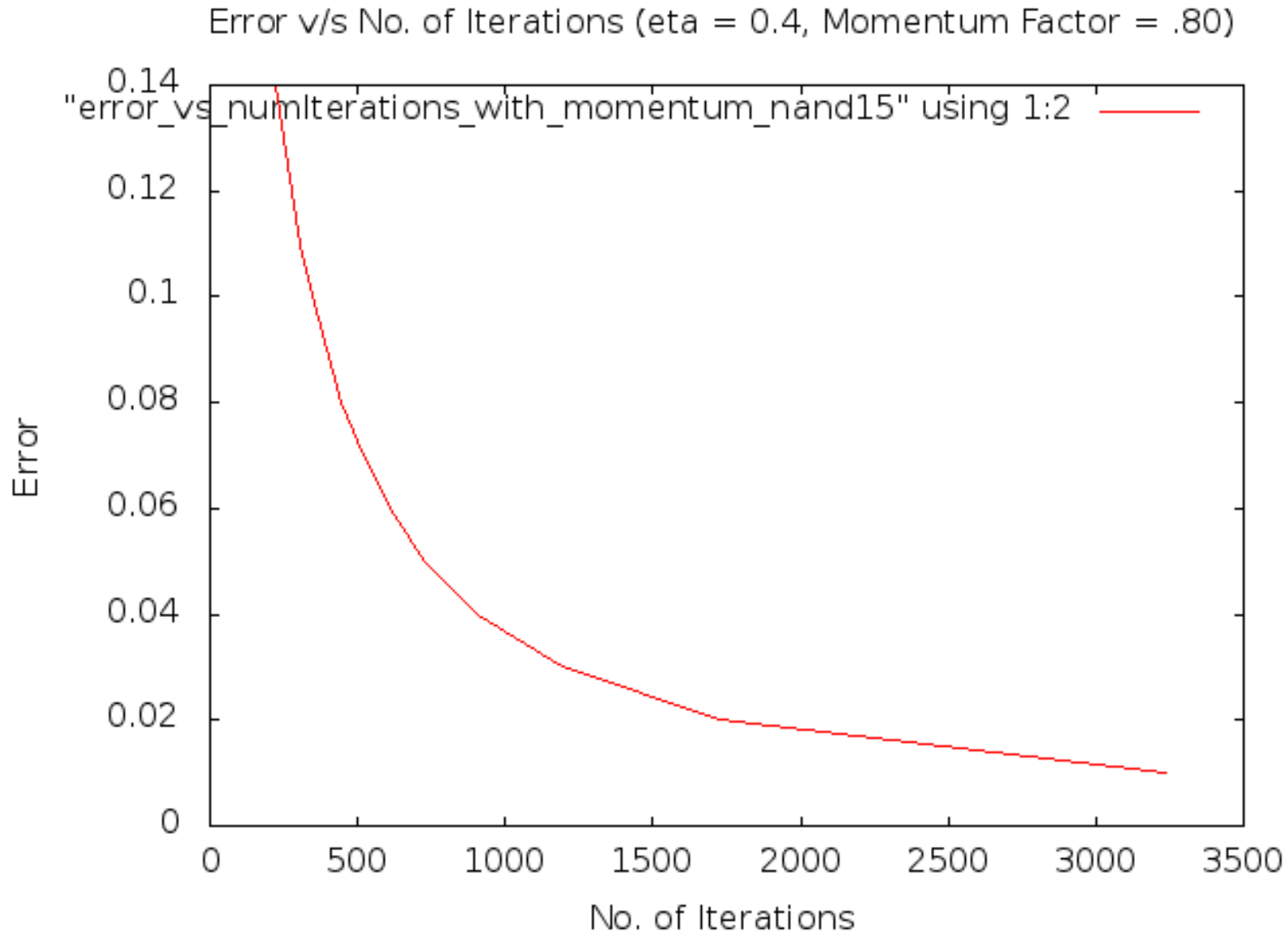




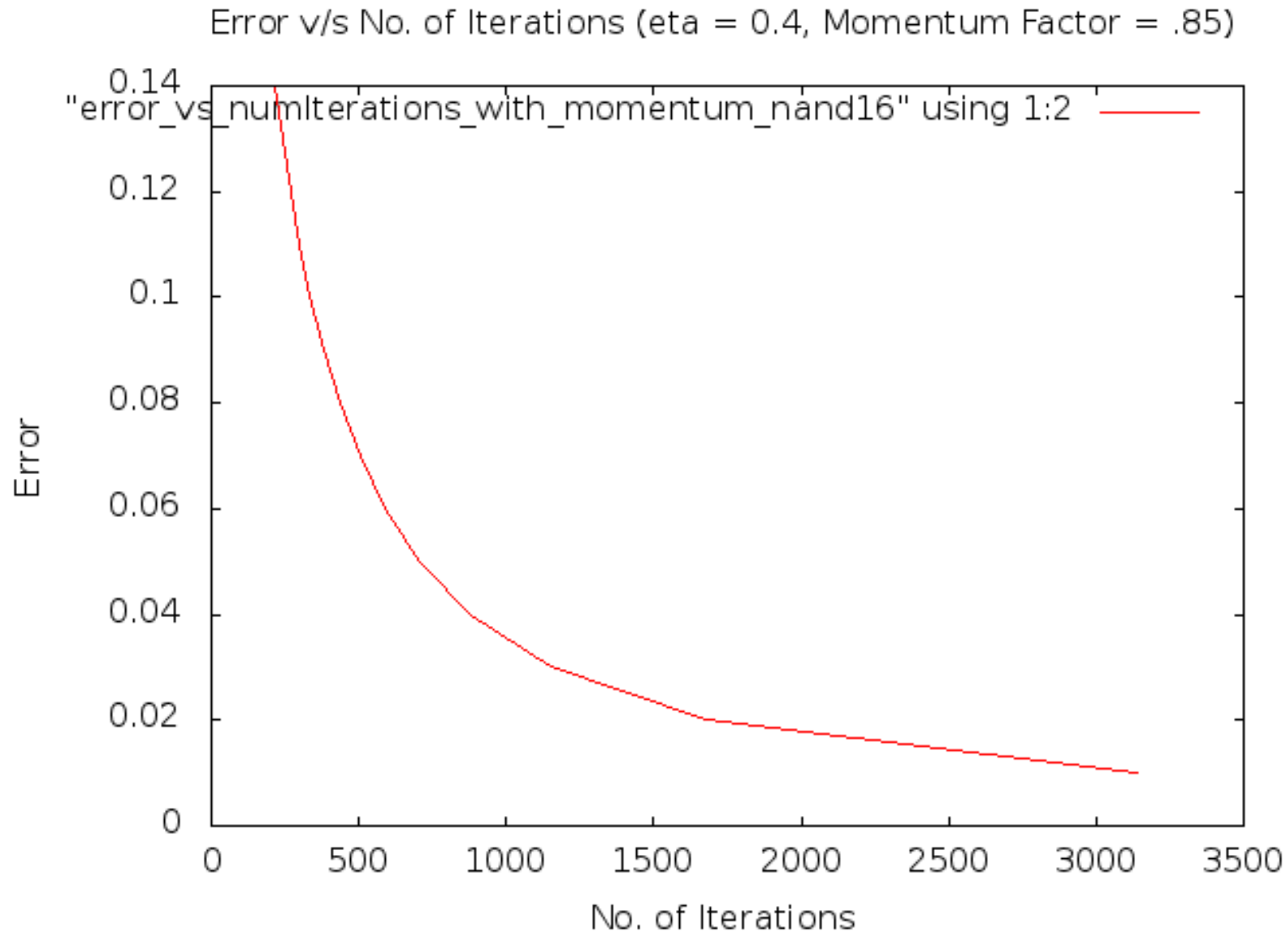
# Effect of Momentum Factor (NAND)



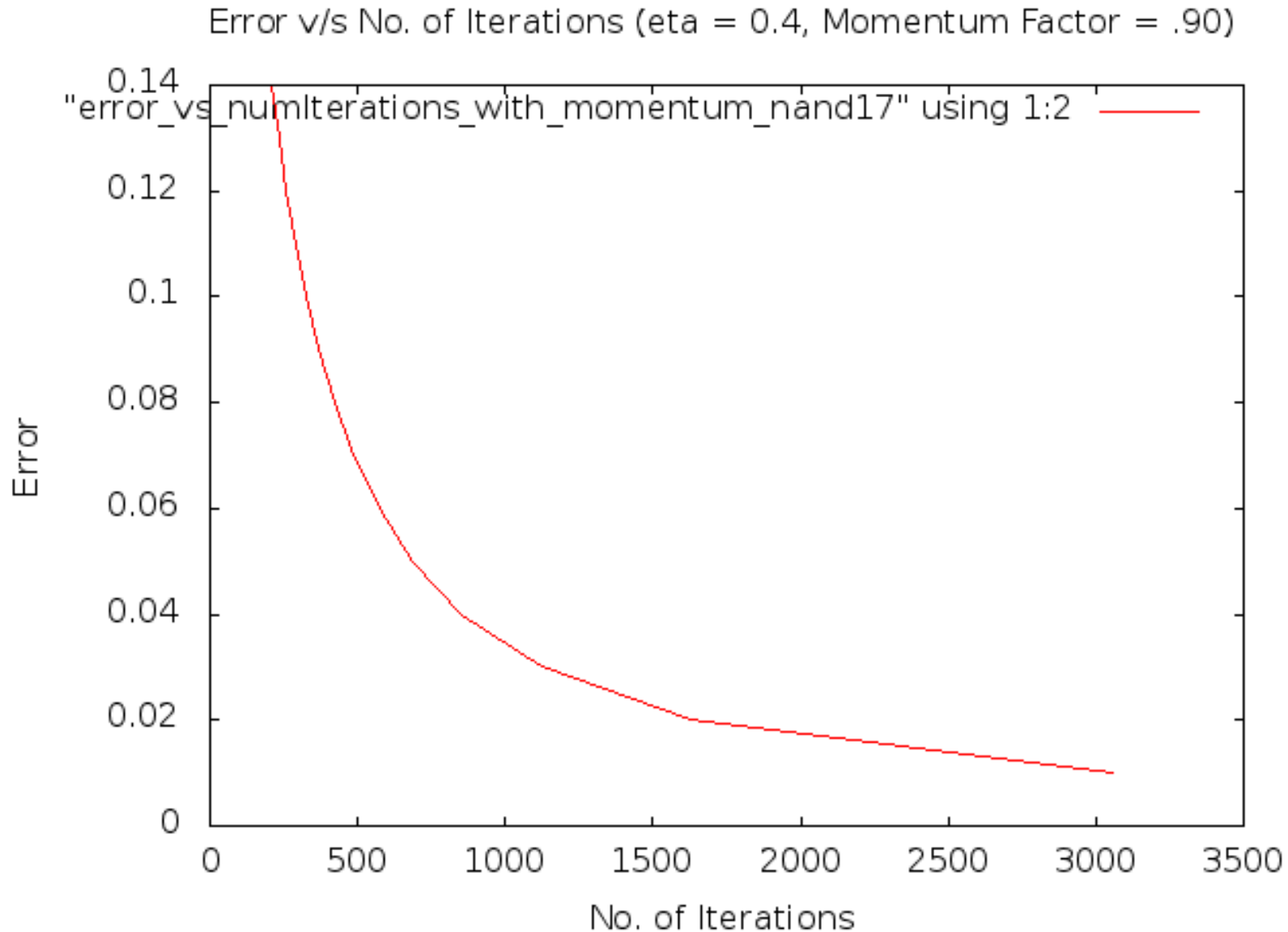
# Effect of Momentum Factor (NAND)



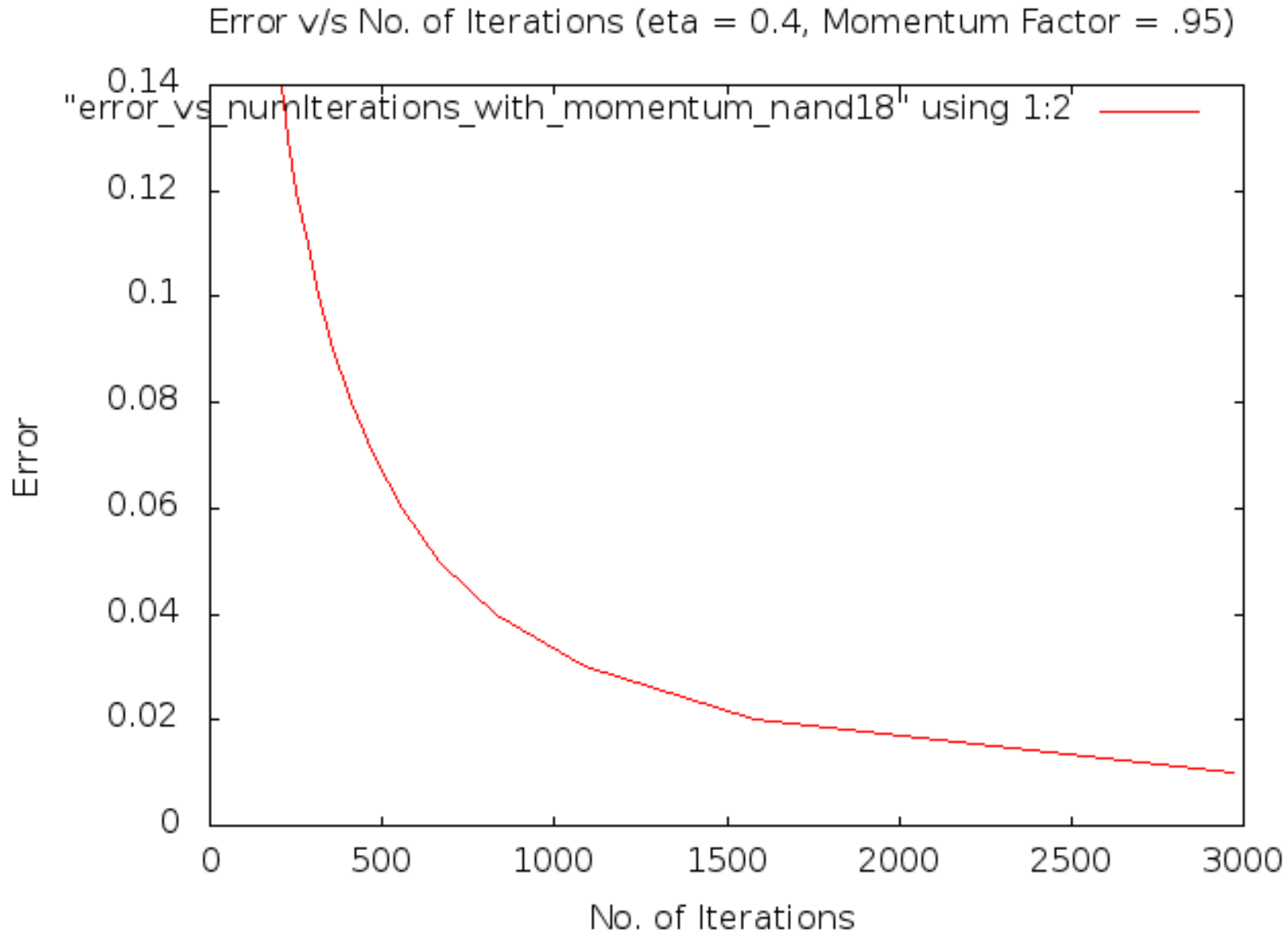
# Effect of Momentum Factor (NAND)



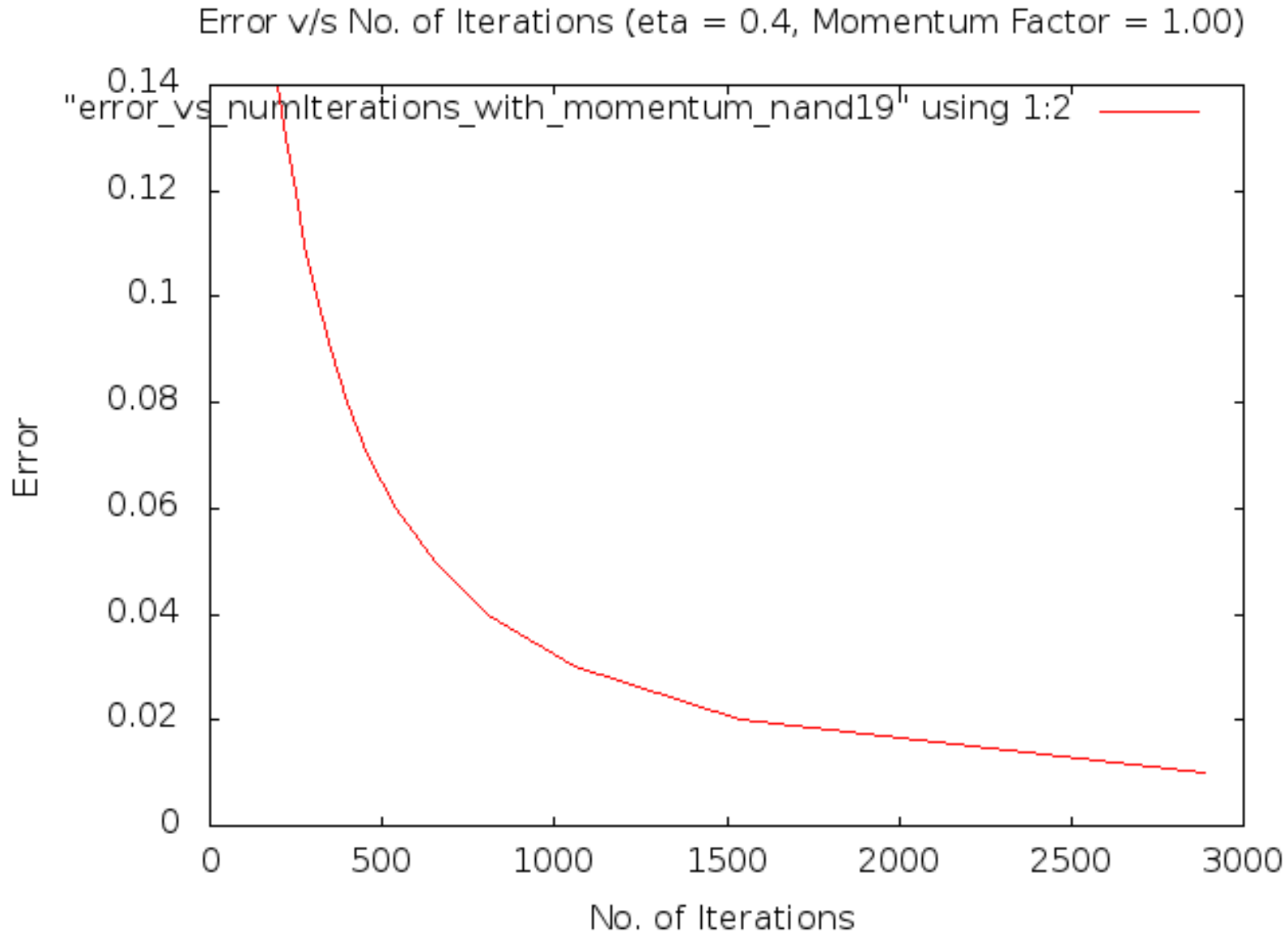
# Effect of Momentum Factor (NAND)



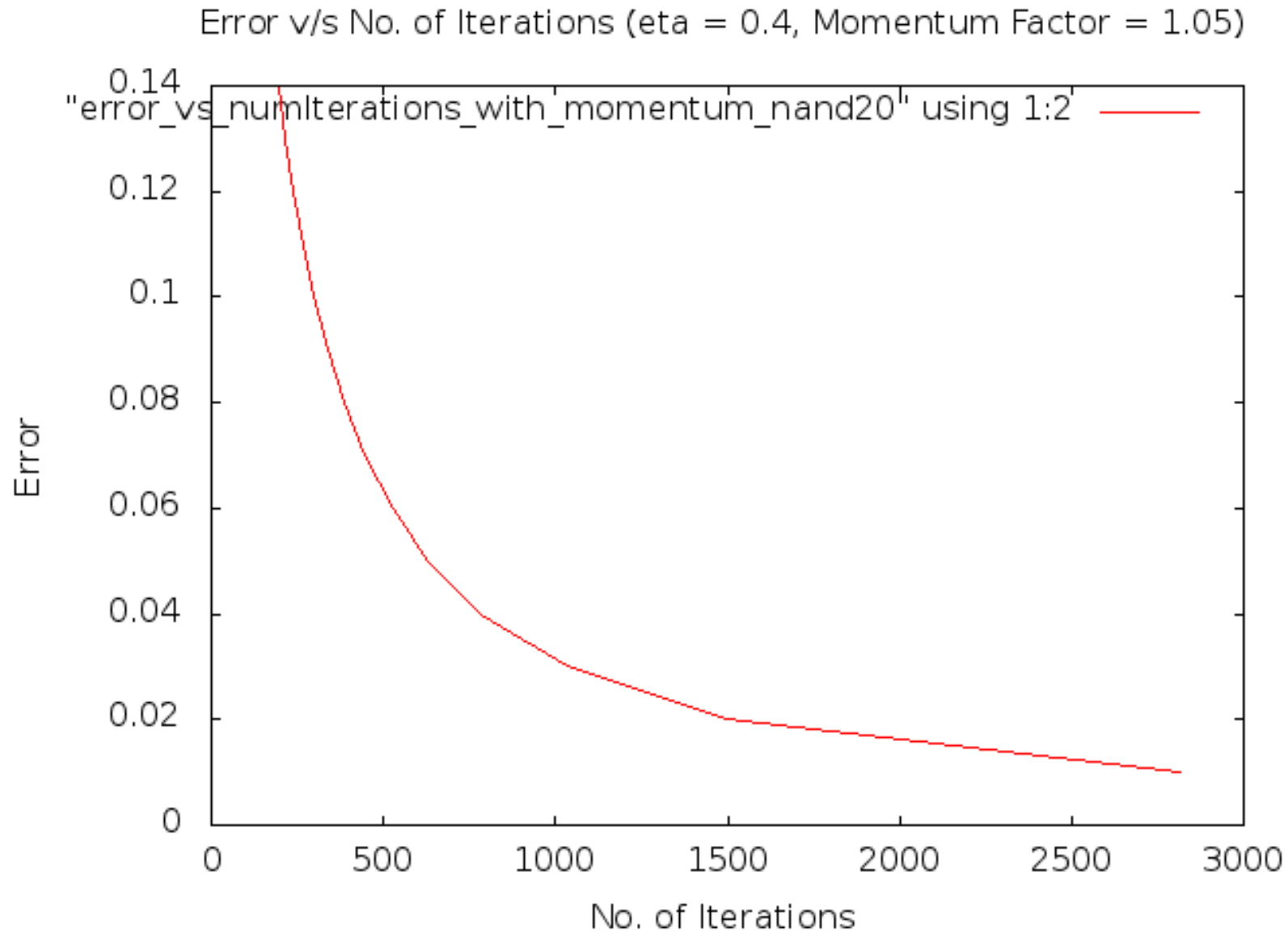
# Effect of Momentum Factor (NAND)



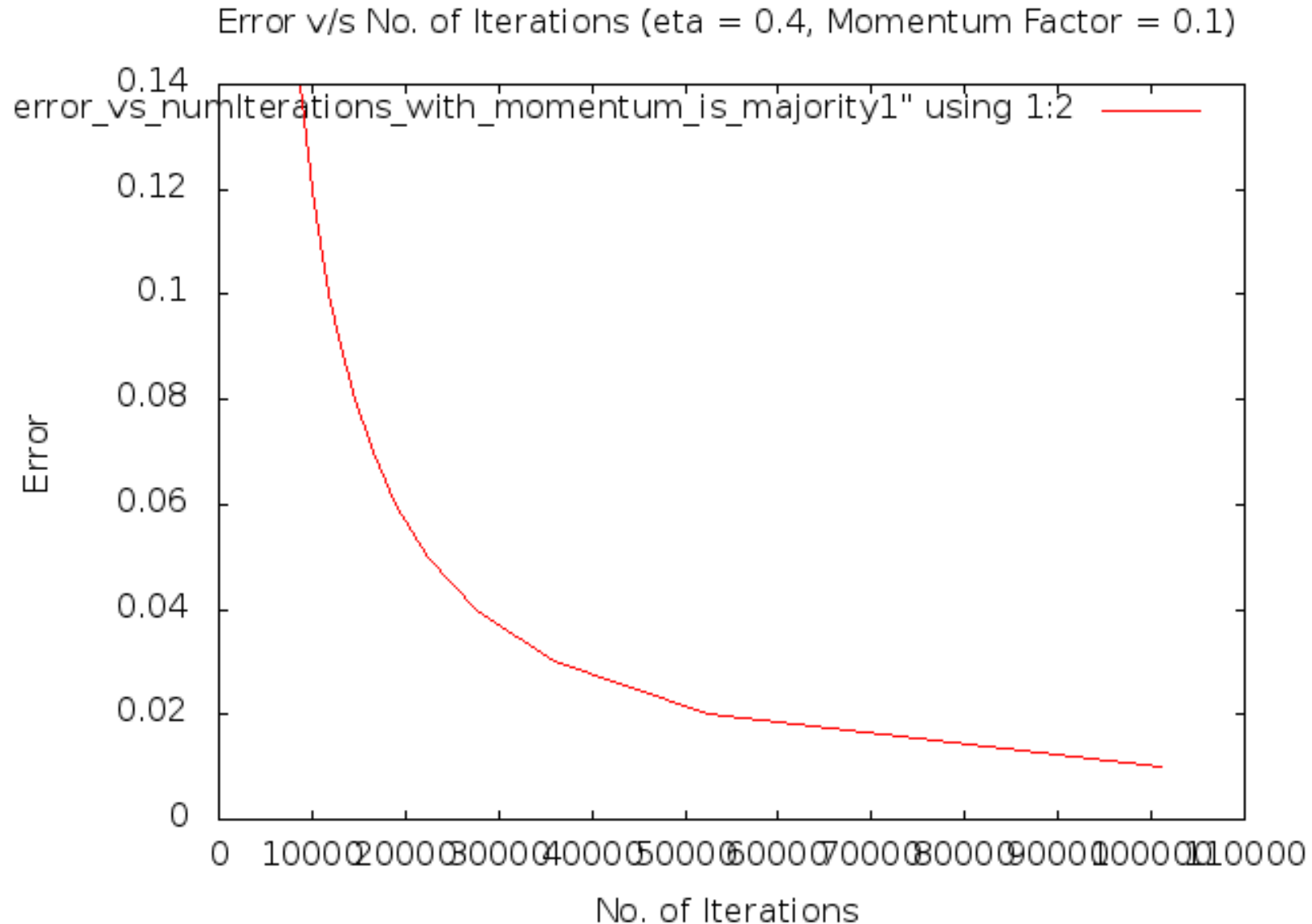
# Effect of Momentum Factor (NAND)



# Effect of Momentum Factor (NAND)

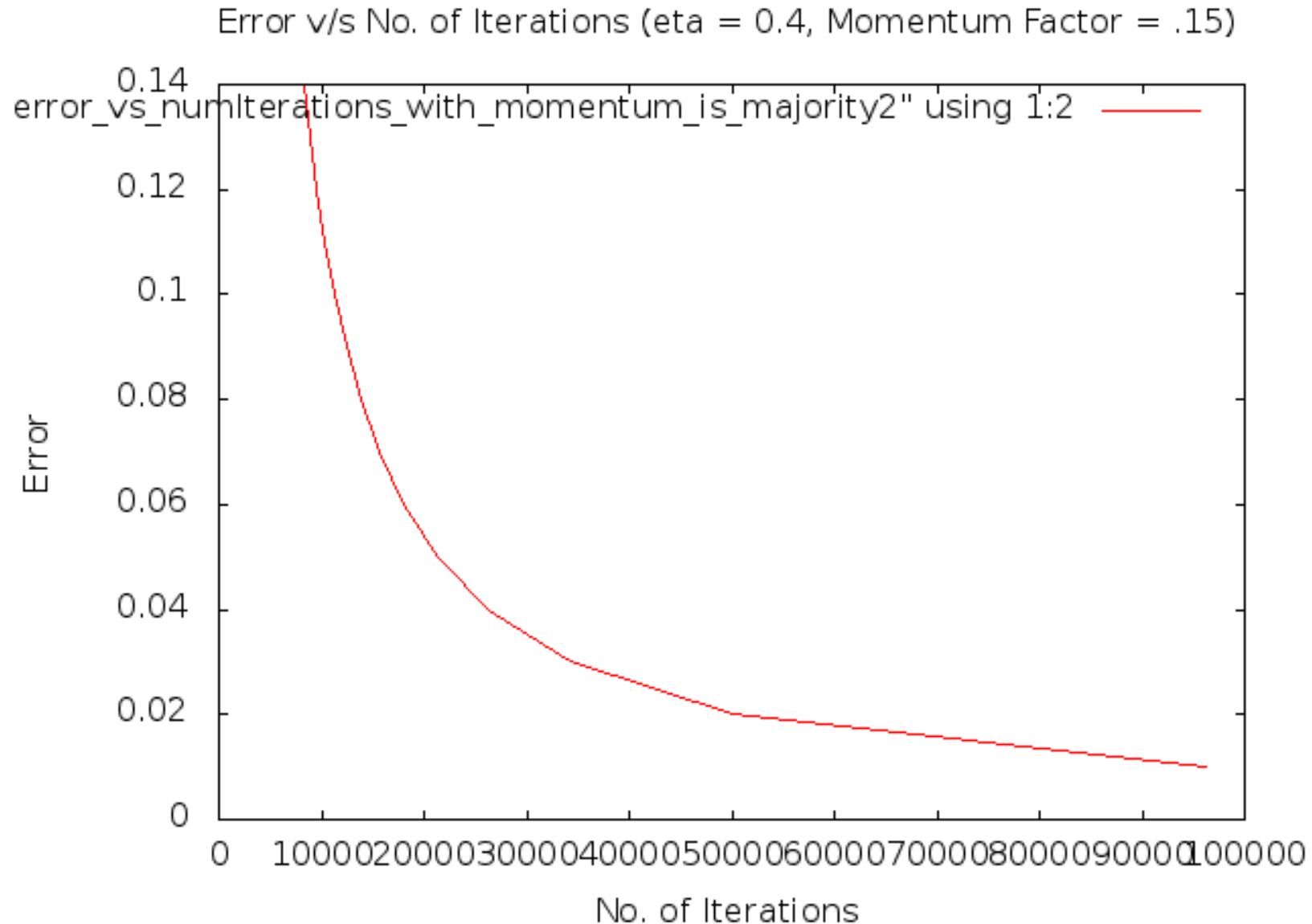


# Effect of Momentum Factor (Majority)

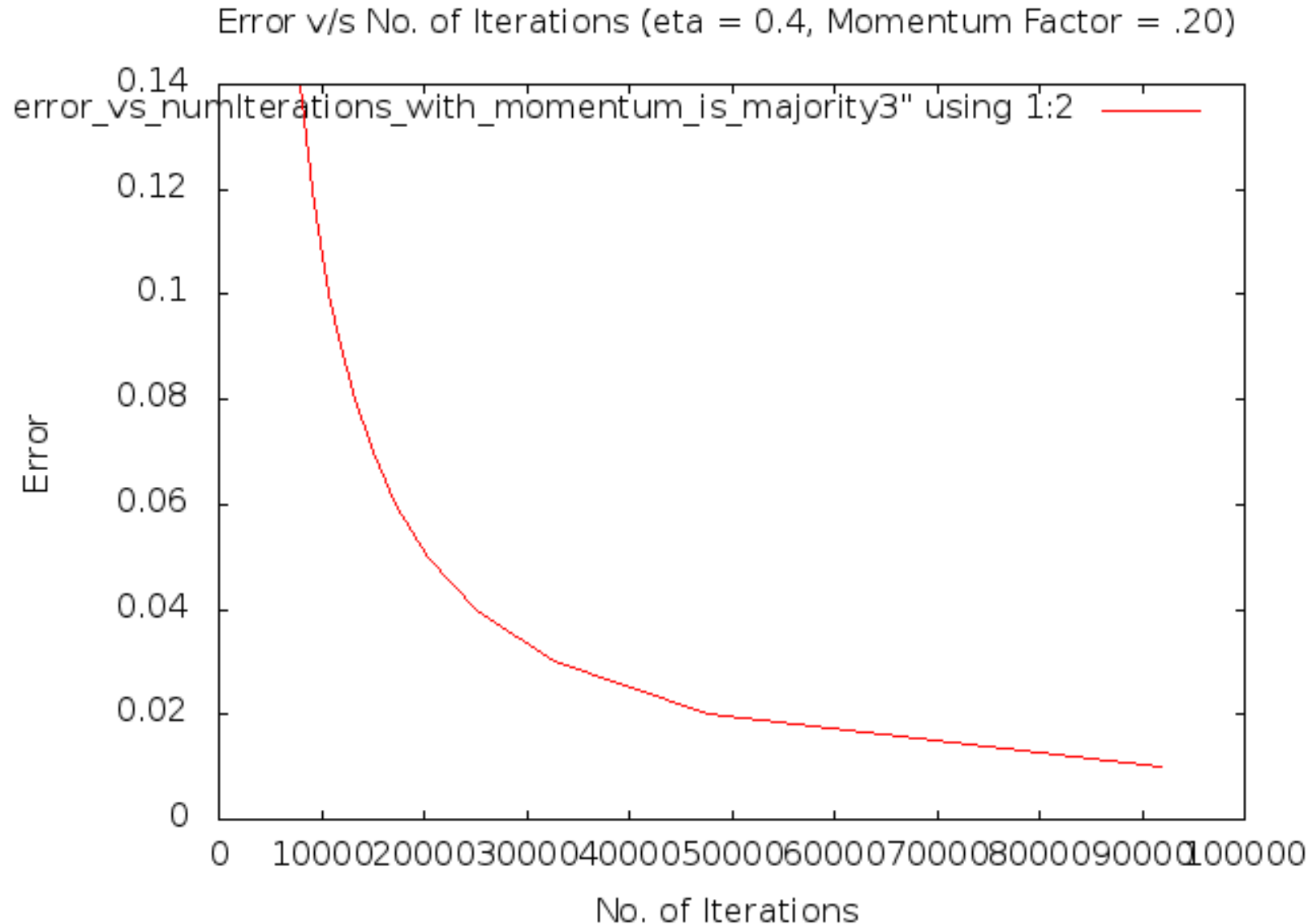




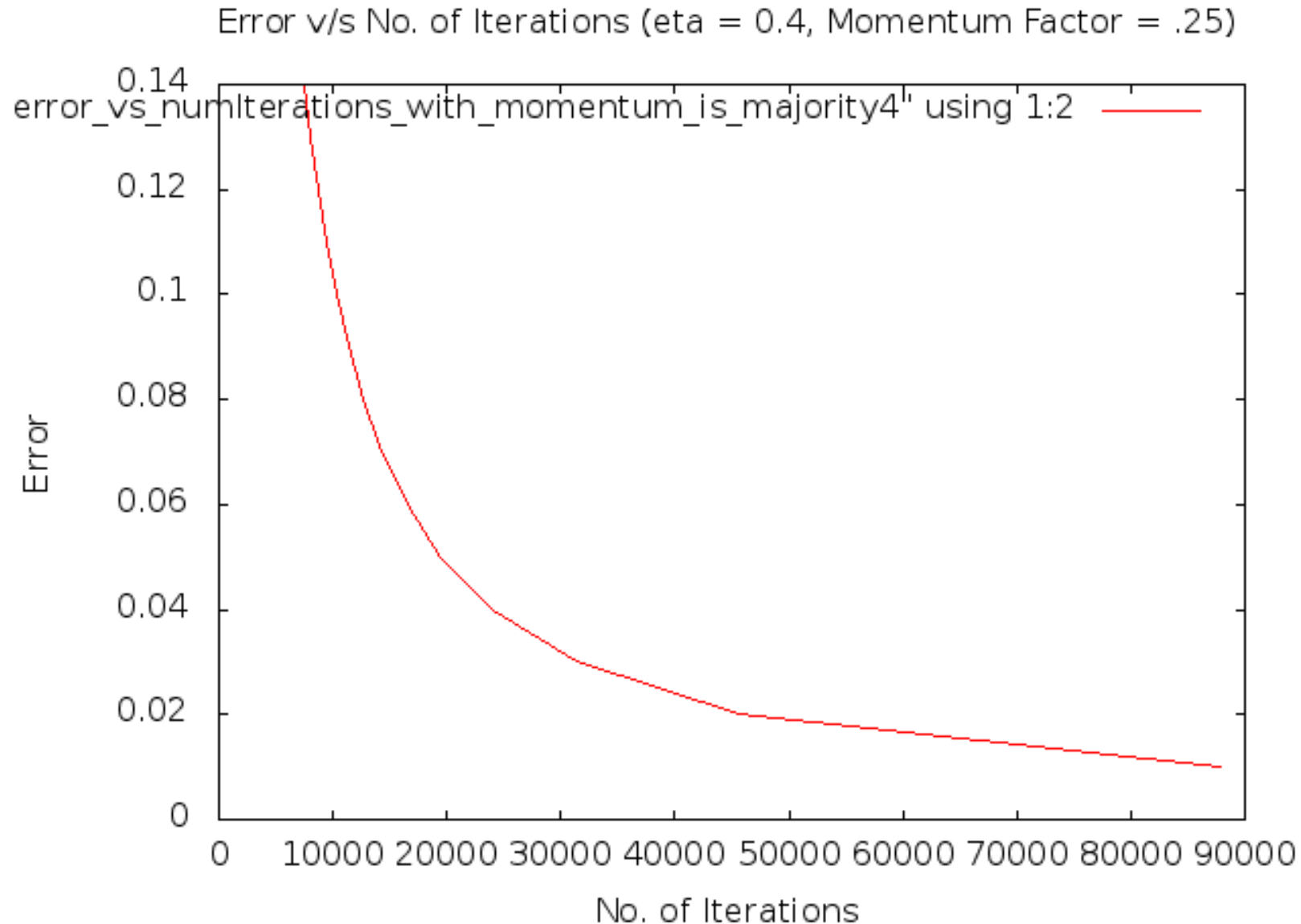
# Effect of Momentum Factor (Majority)



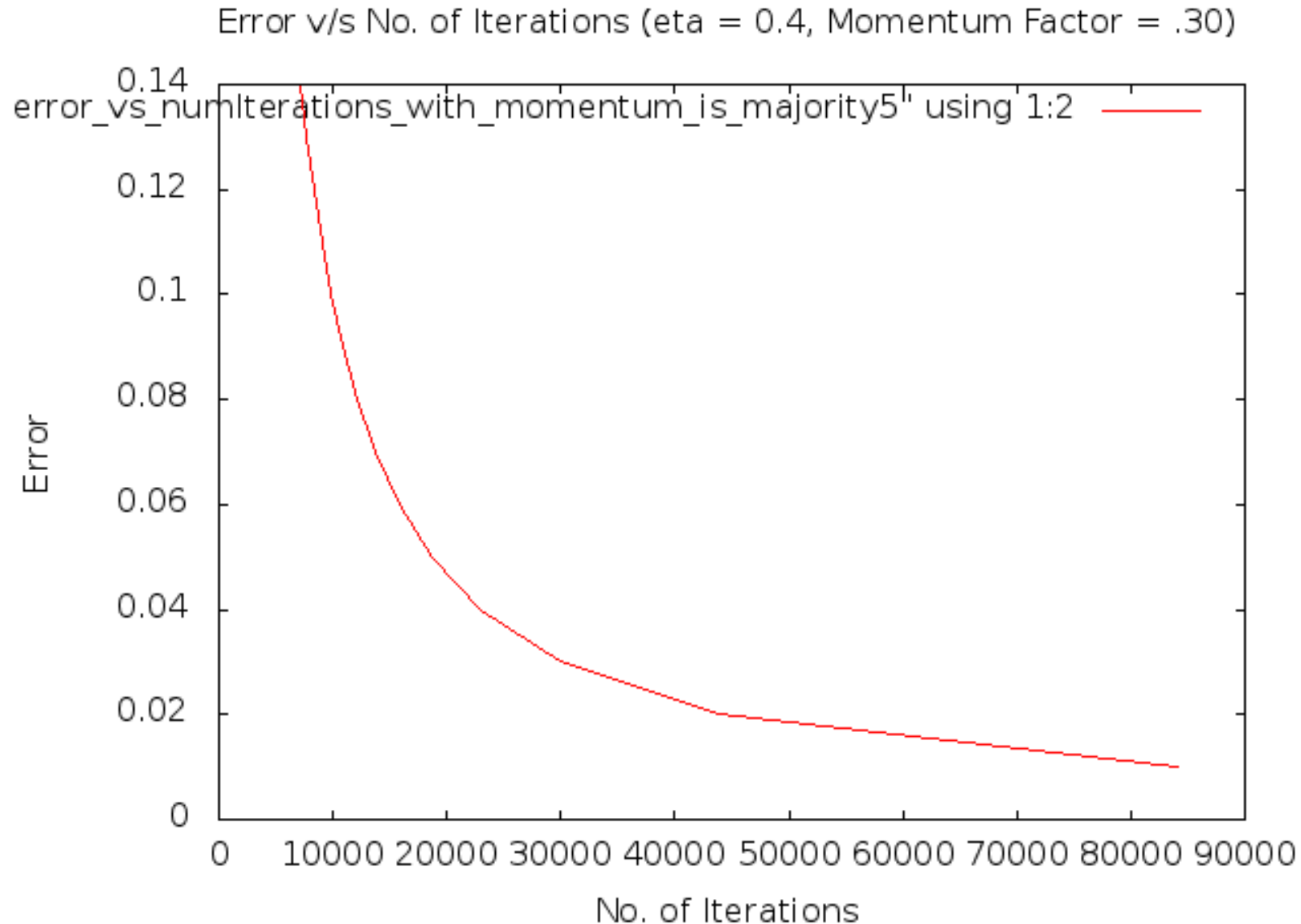
# Effect of Momentum Factor (Majority)



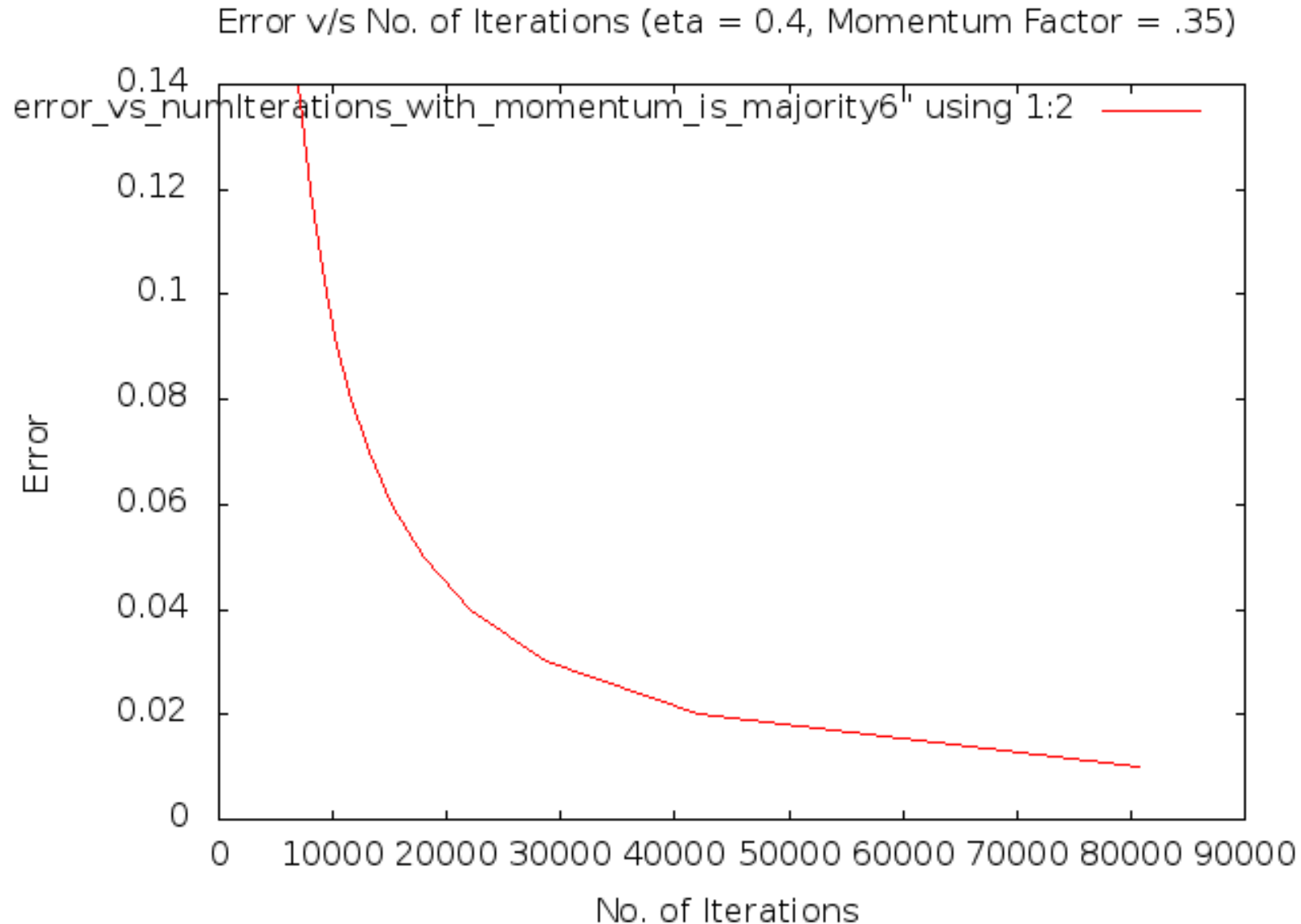
# Effect of Momentum Factor (Majority)



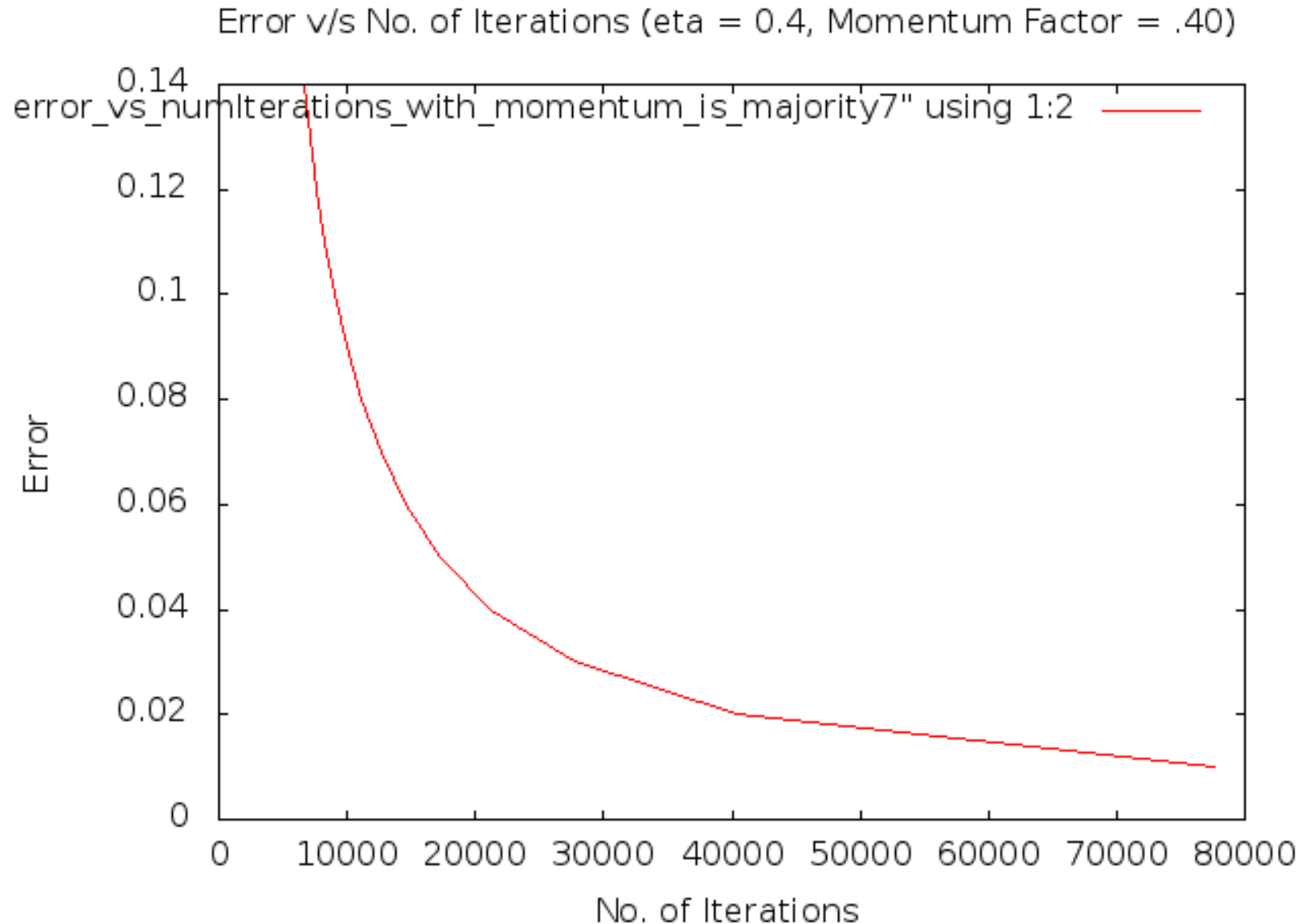
# Effect of Momentum Factor (Majority)



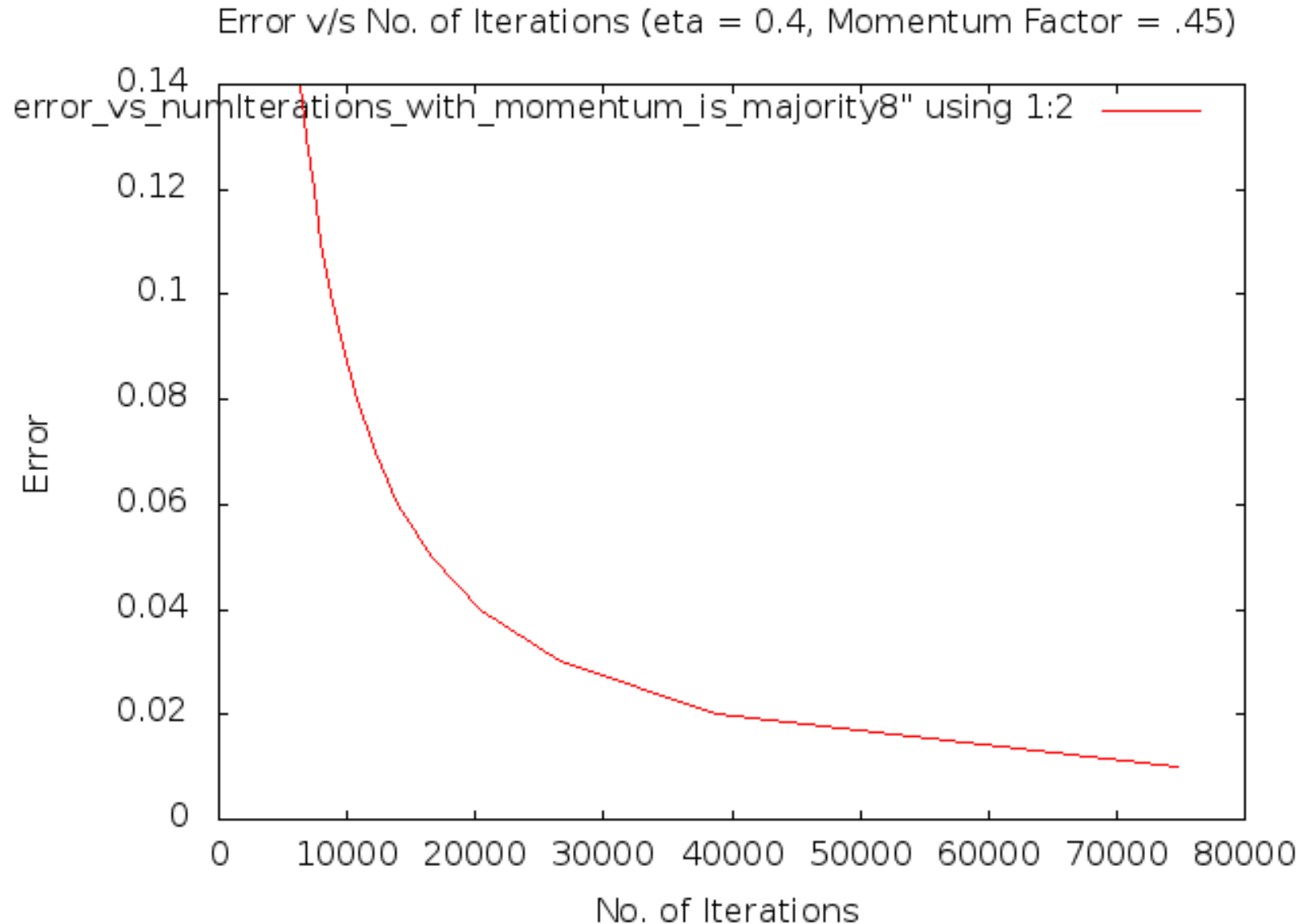
# Effect of Momentum Factor (Majority)



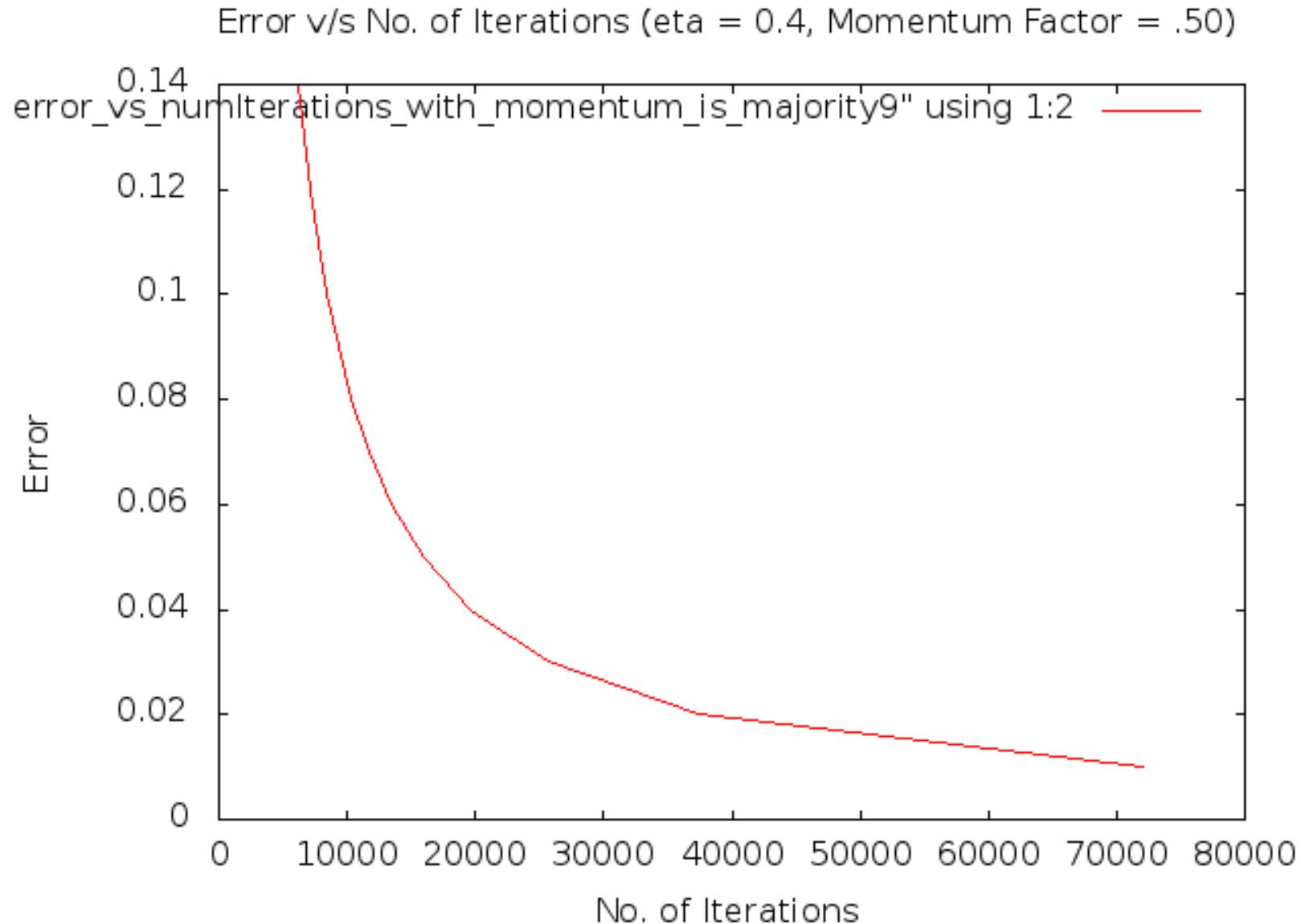
# Effect of Momentum Factor (Majority)



# Effect of Momentum Factor (Majority)

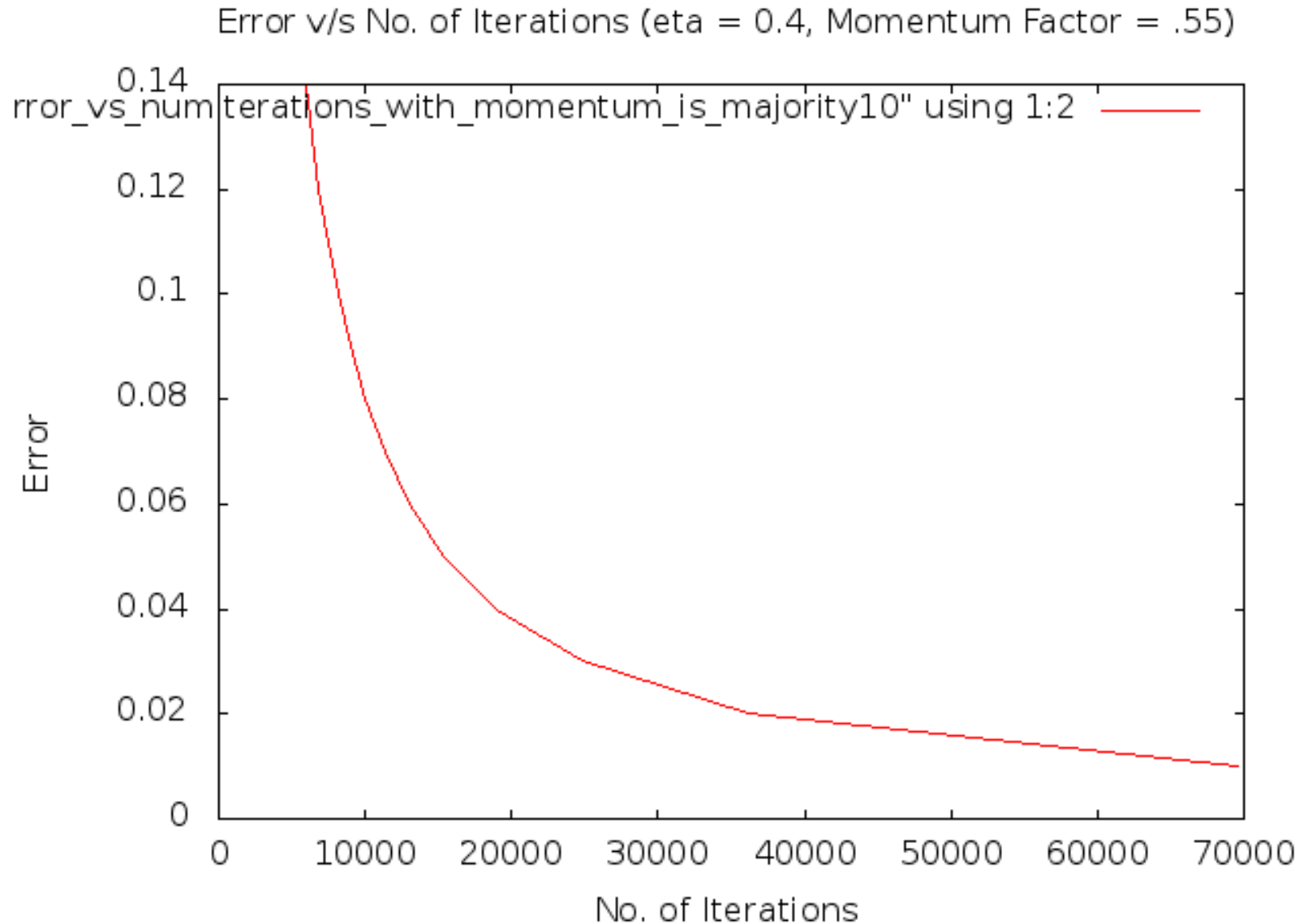


# Effect of Momentum Factor (Majority)

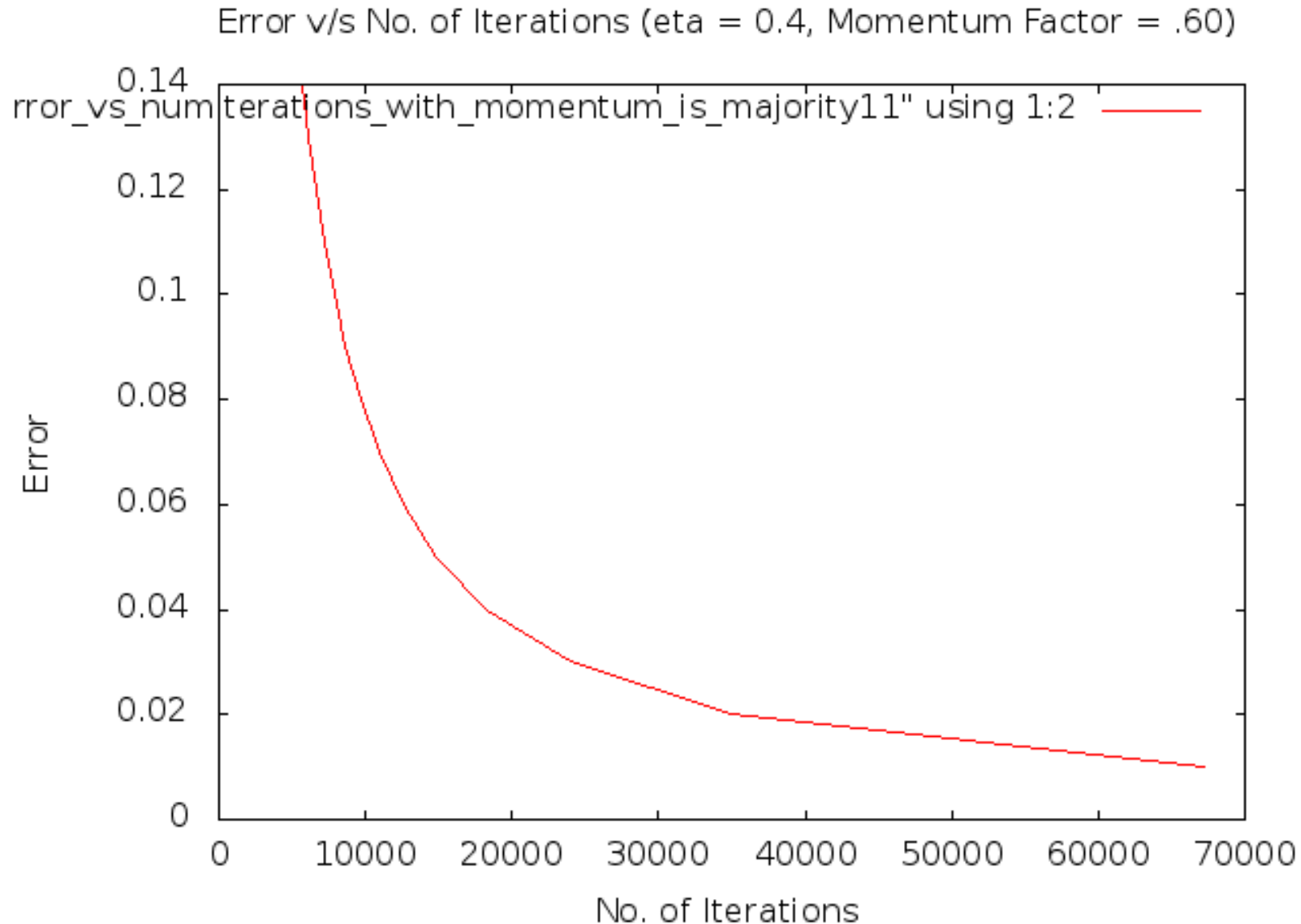




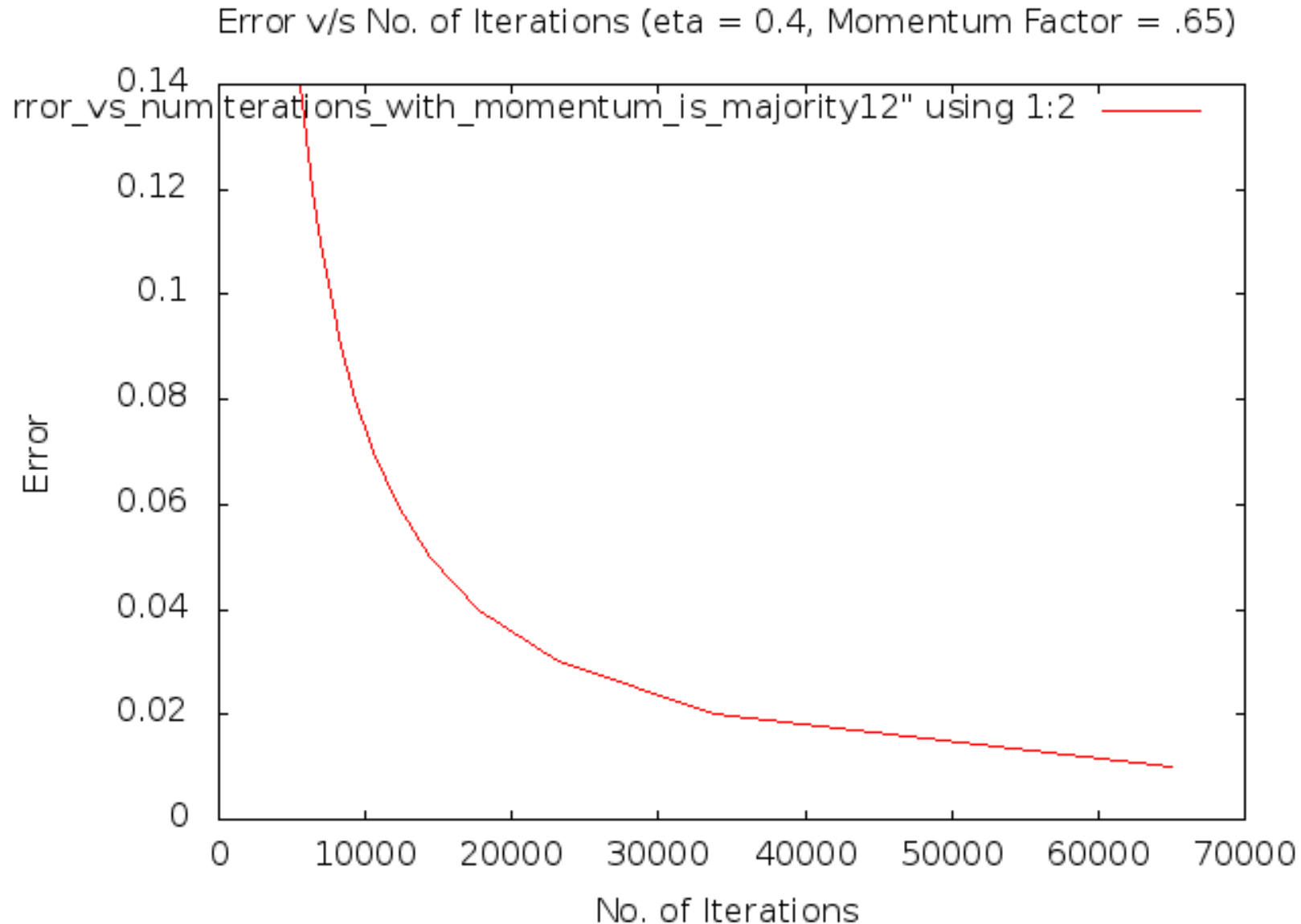
# Effect of Momentum Factor (Majority)



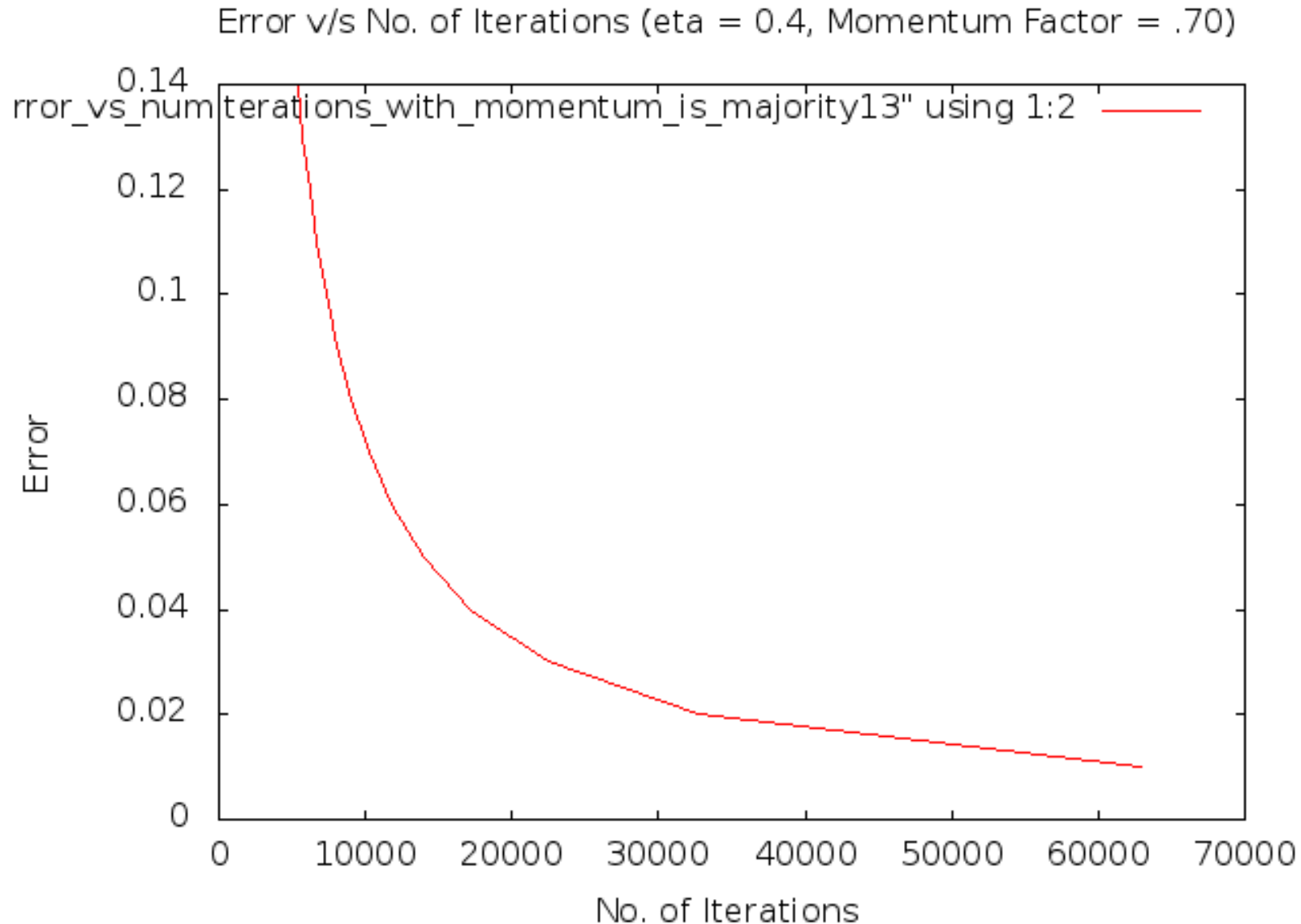
# Effect of Momentum Factor (Majority)



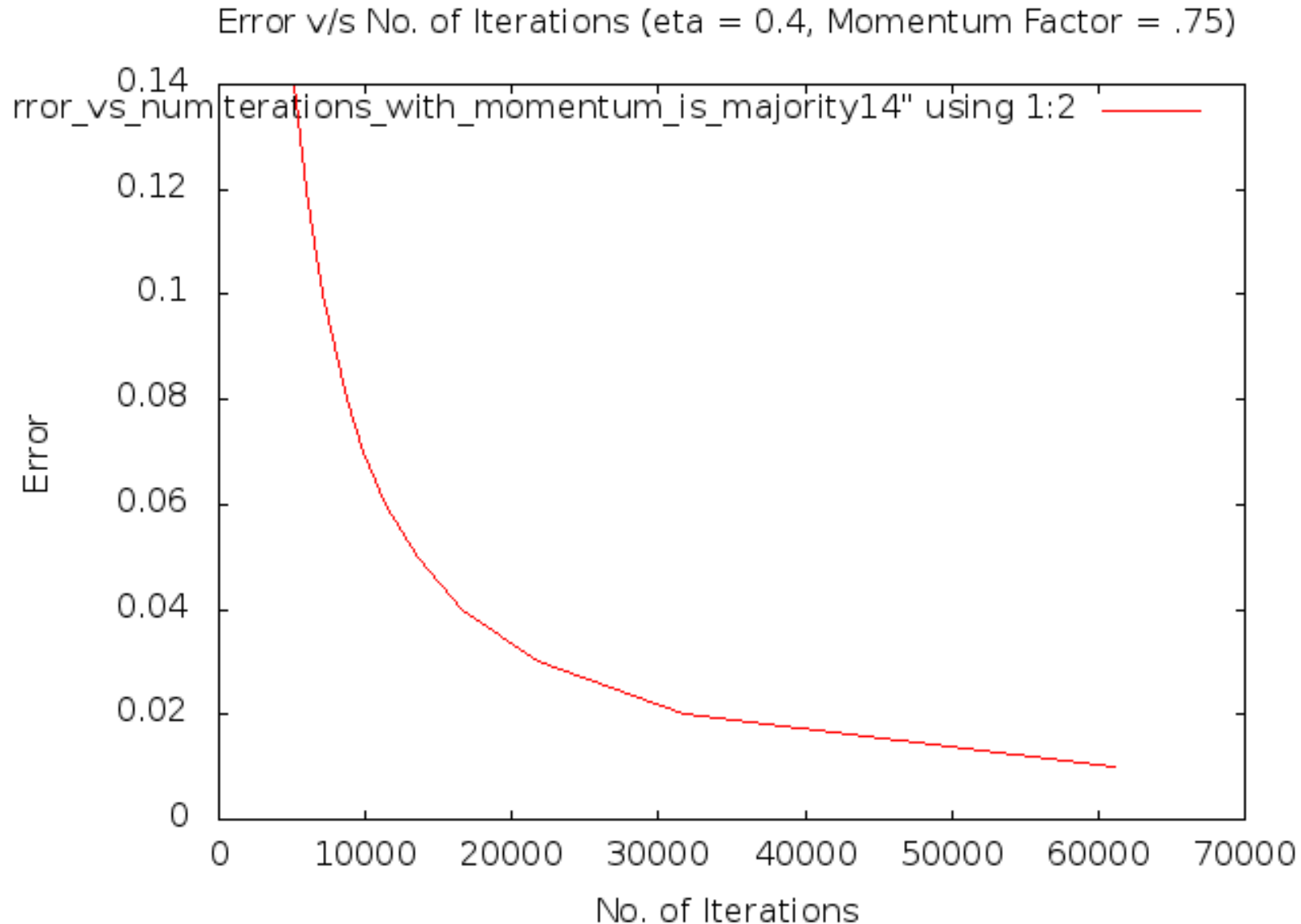
# Effect of Momentum Factor (Majority)



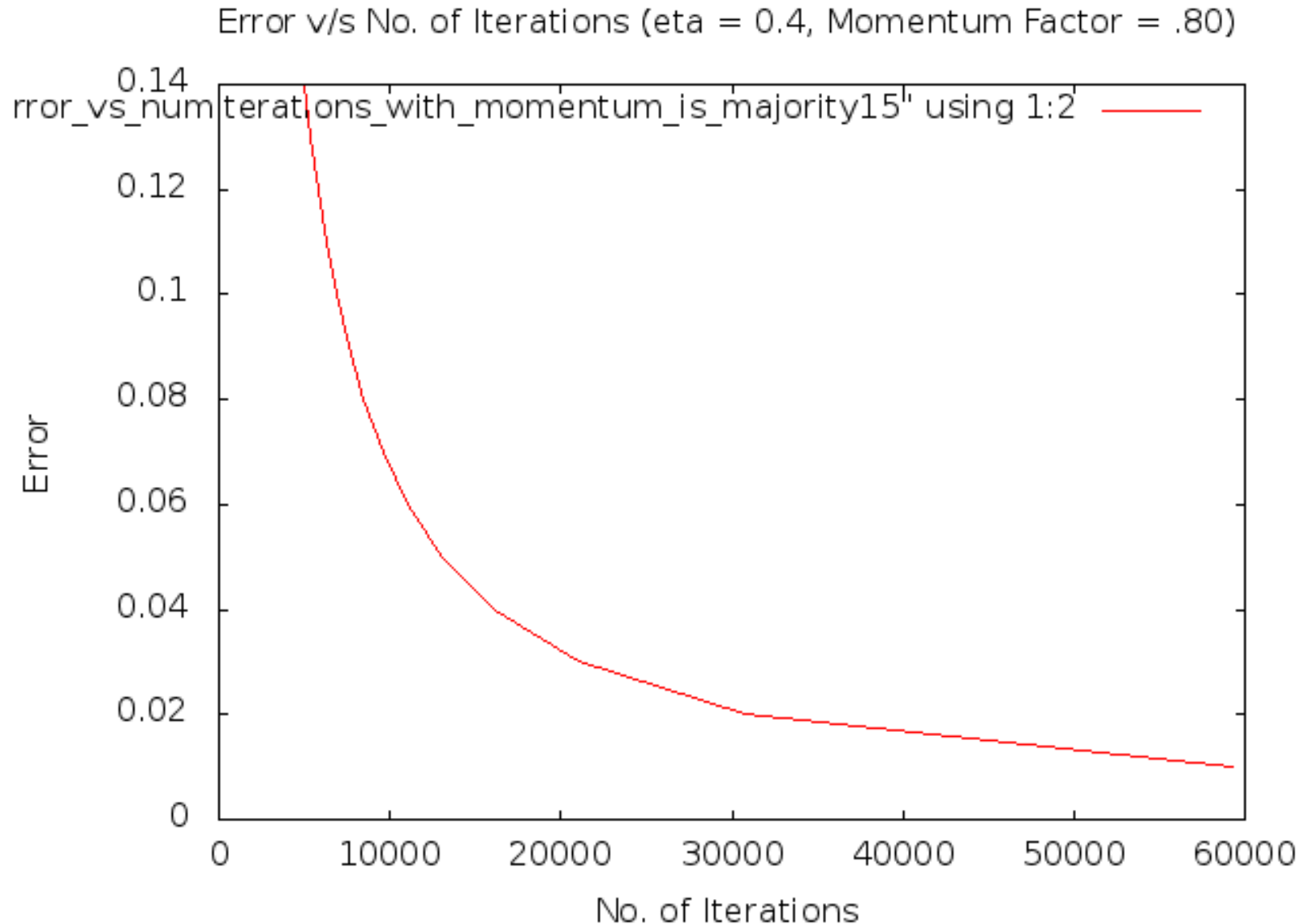
# Effect of Momentum Factor (Majority)



# Effect of Momentum Factor (Majority)



# Effect of Momentum Factor (Majority)



# Effect of Momentum Factor (Majority)

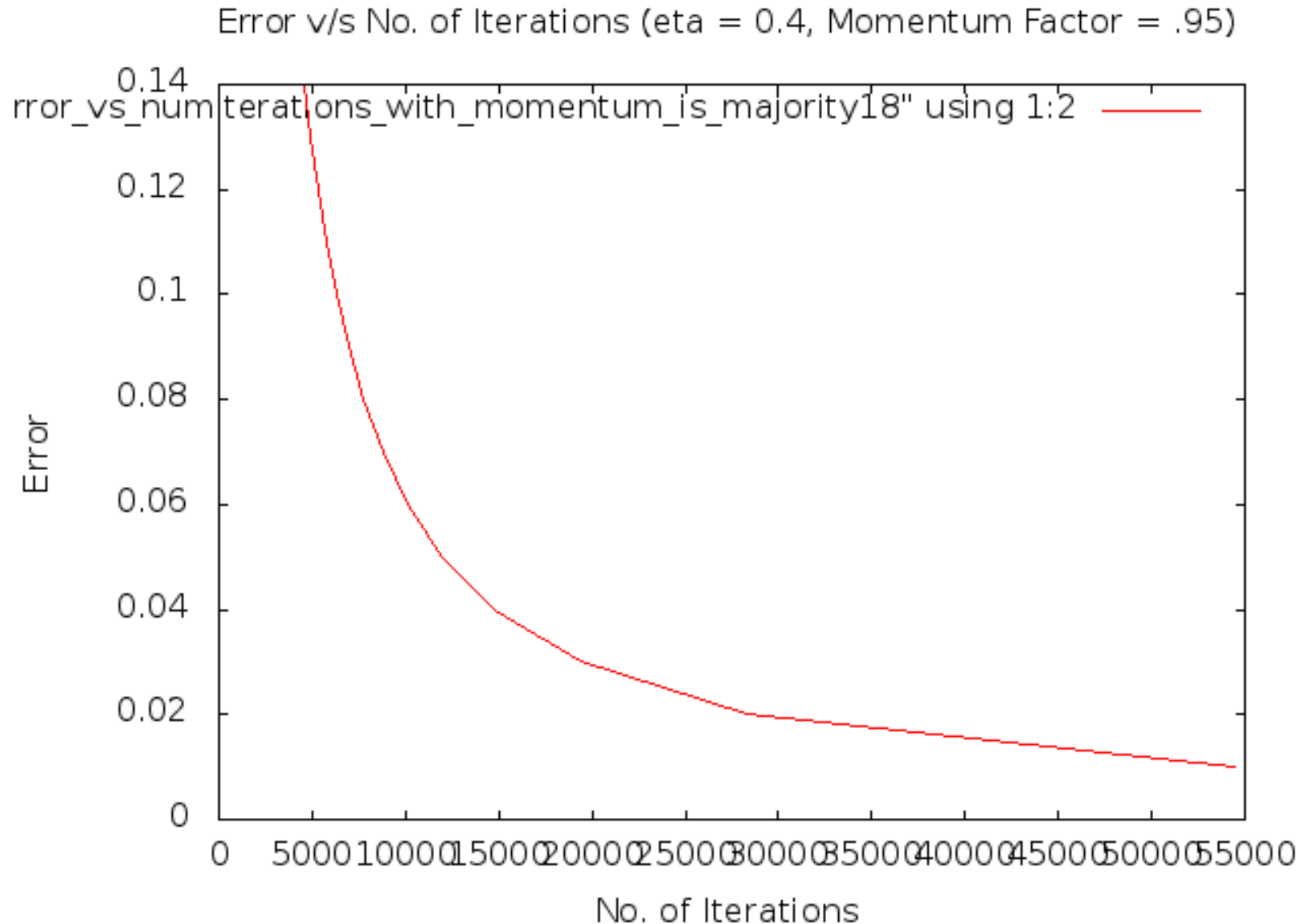


# Effect of Momentum Factor (Majority)

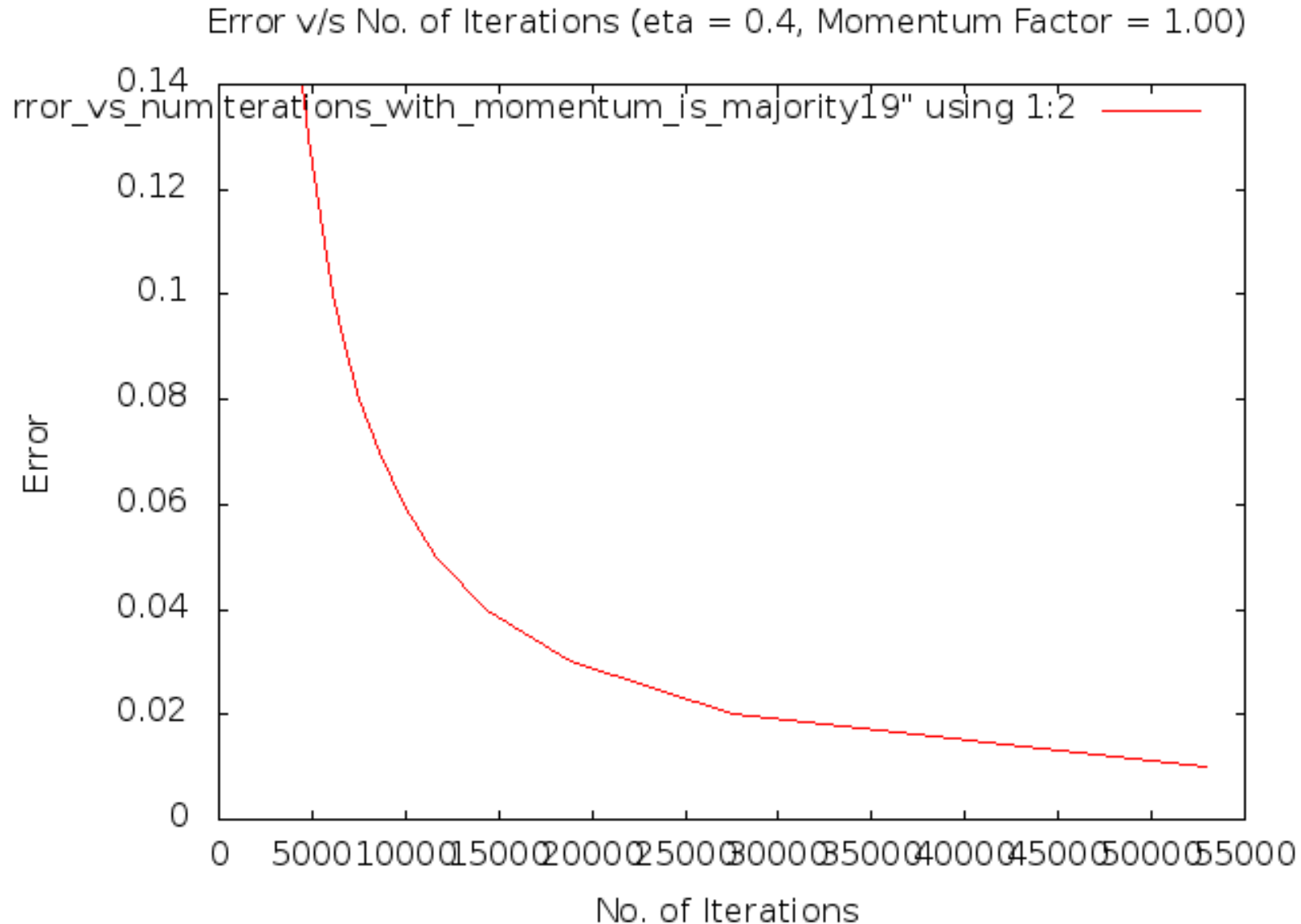




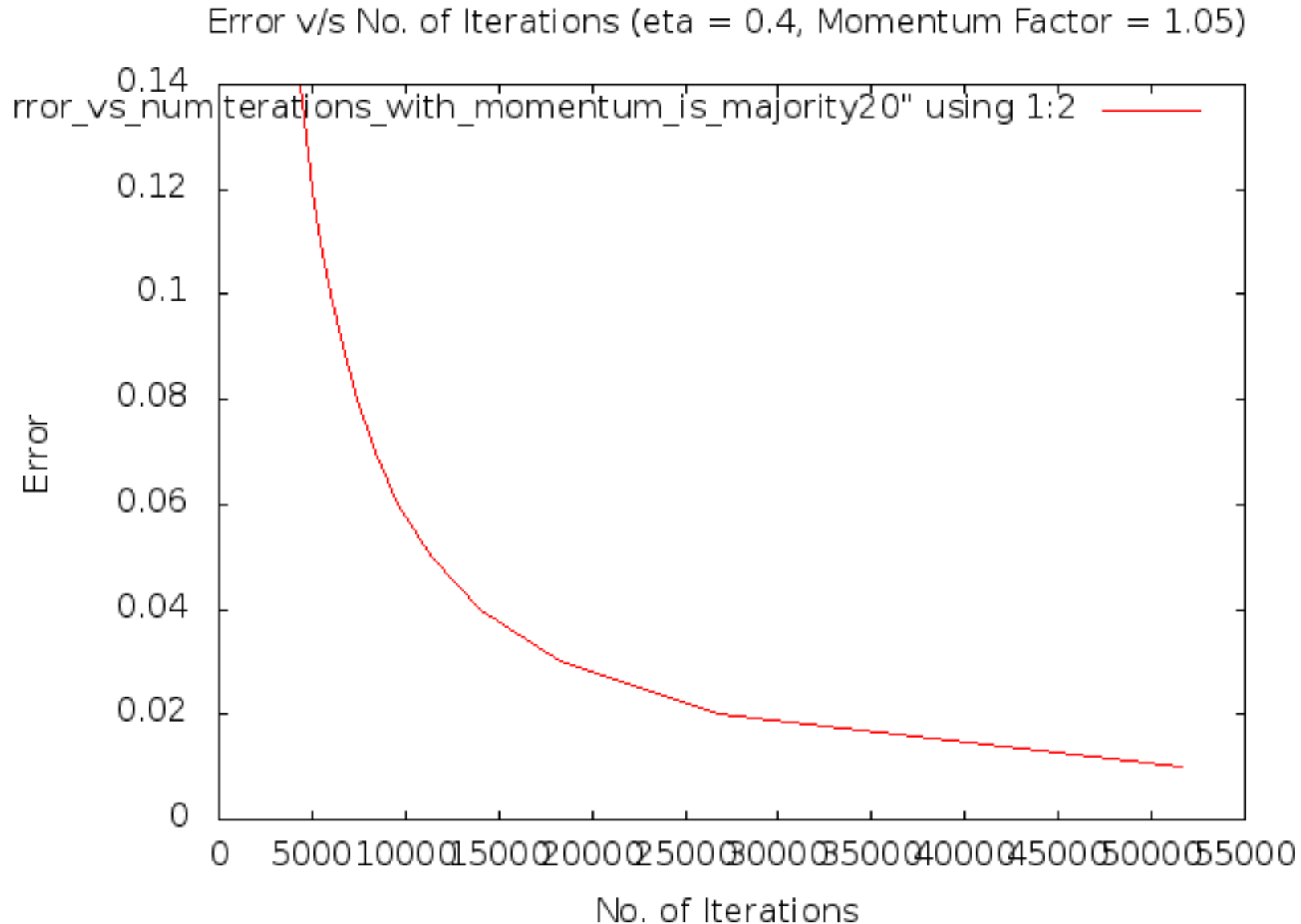
# Effect of Momentum Factor (Majority)



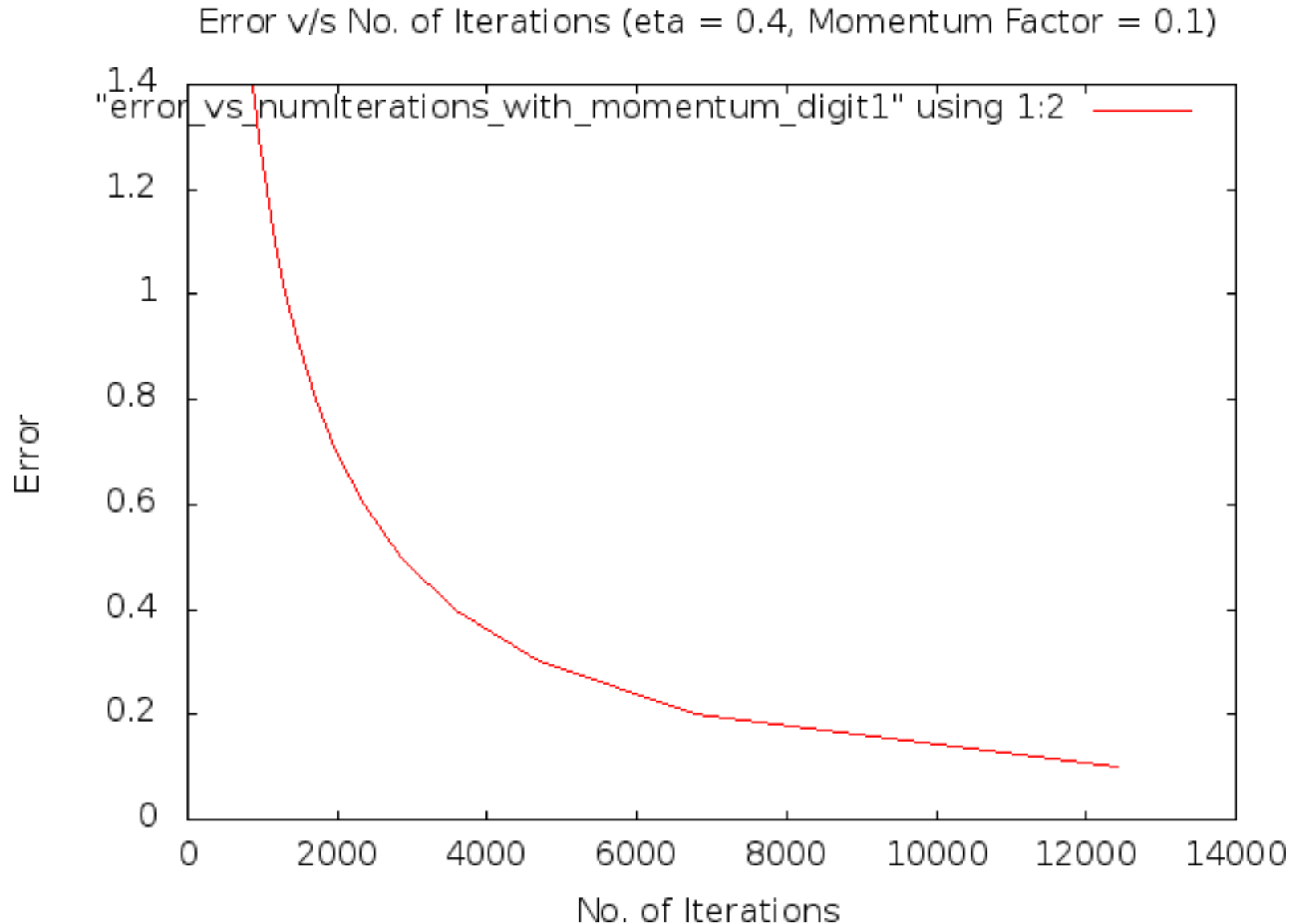
# Effect of Momentum Factor (Majority)



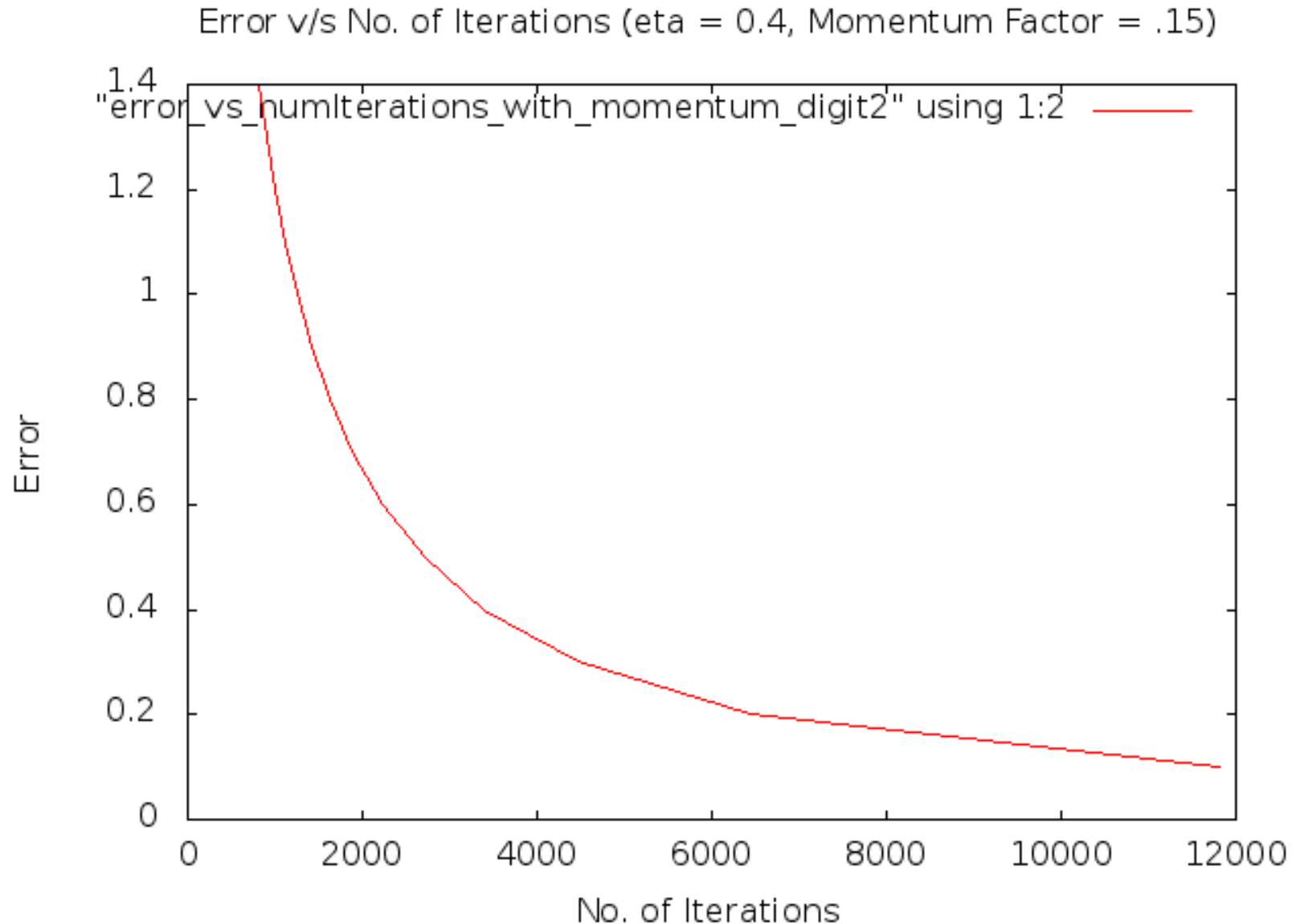
# Effect of Momentum Factor (Majority)



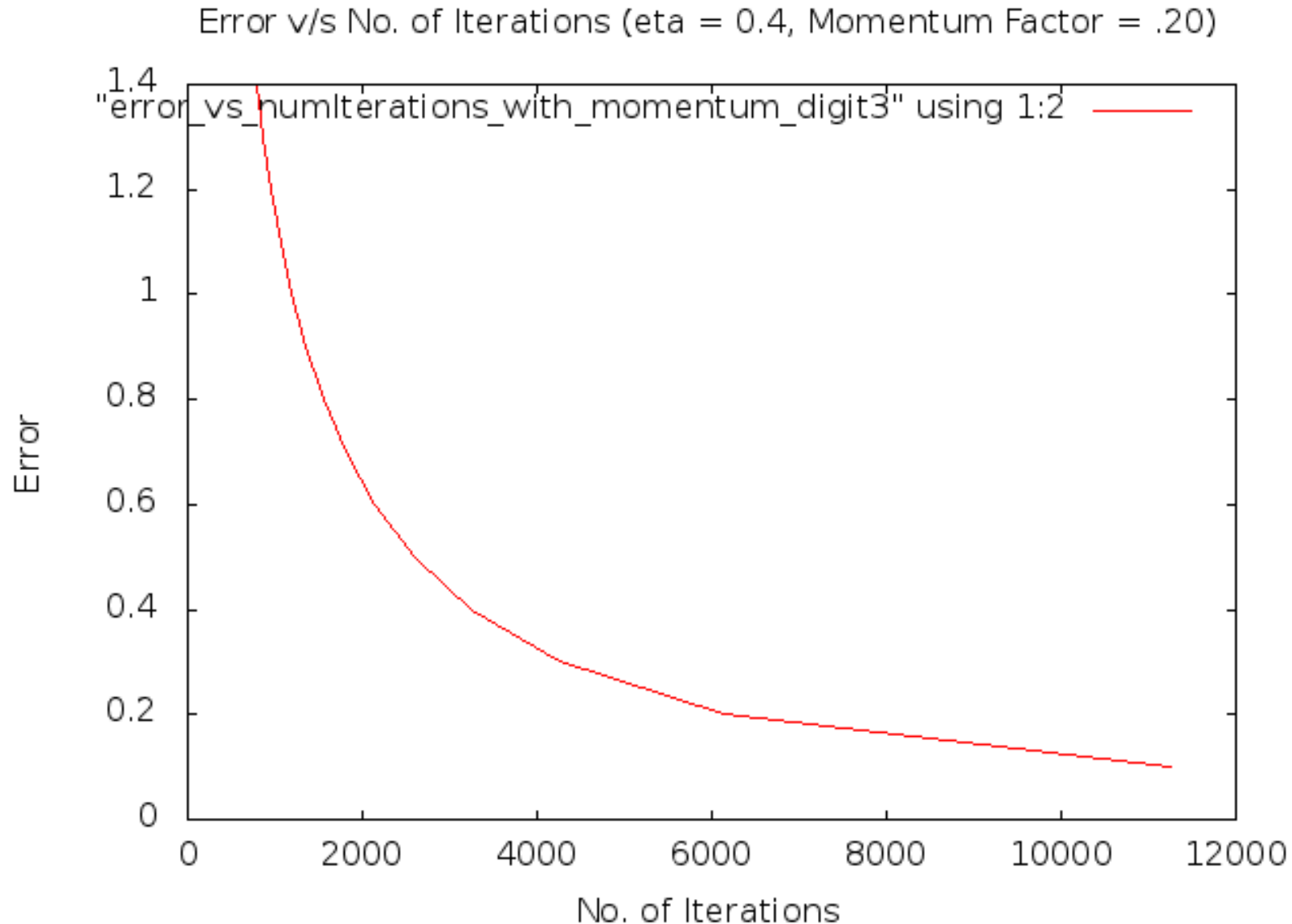
# Effect of Momentum Factor (Digit Recogniser)



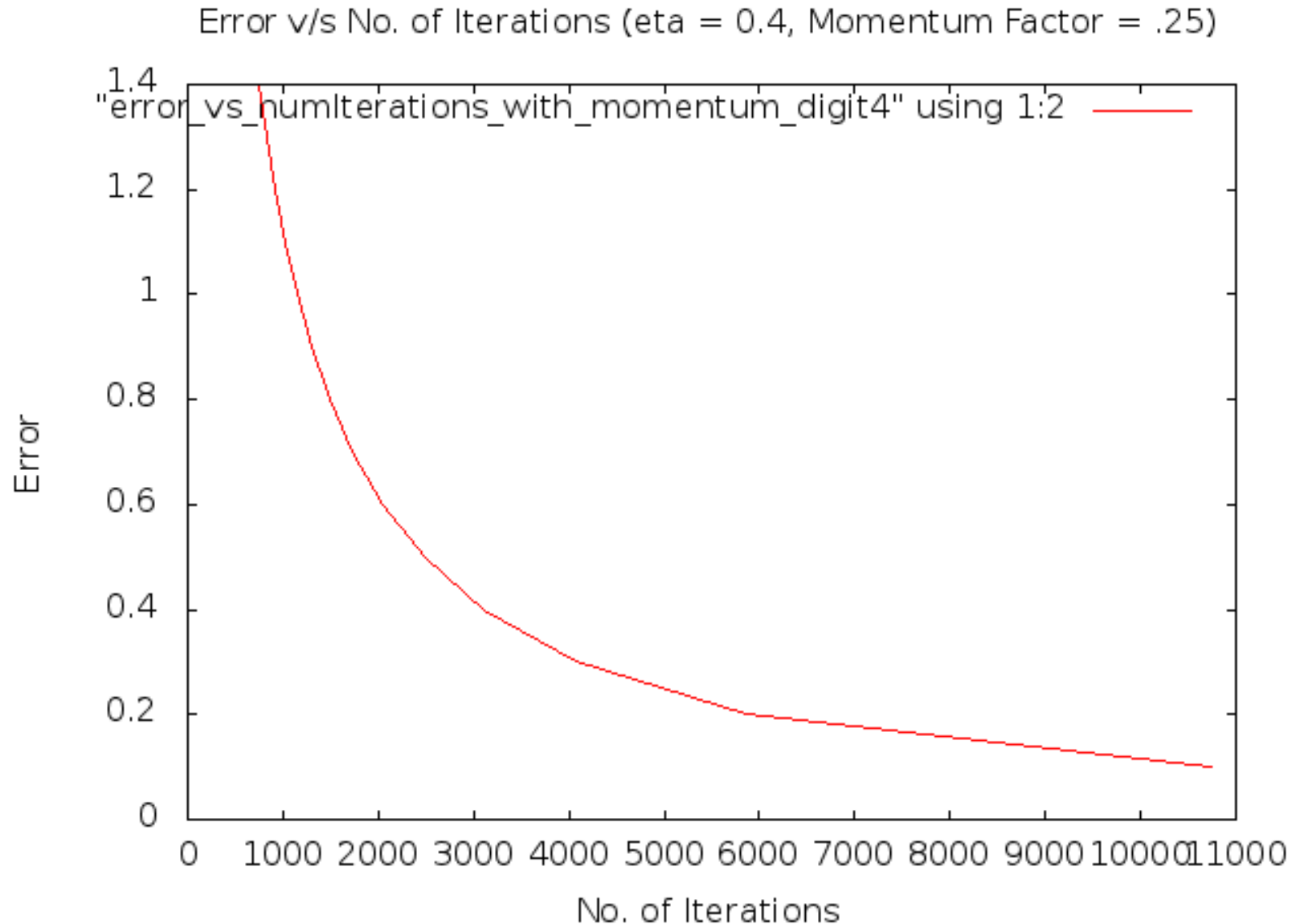
# Effect of Momentum Factor (Digit Recogniser)



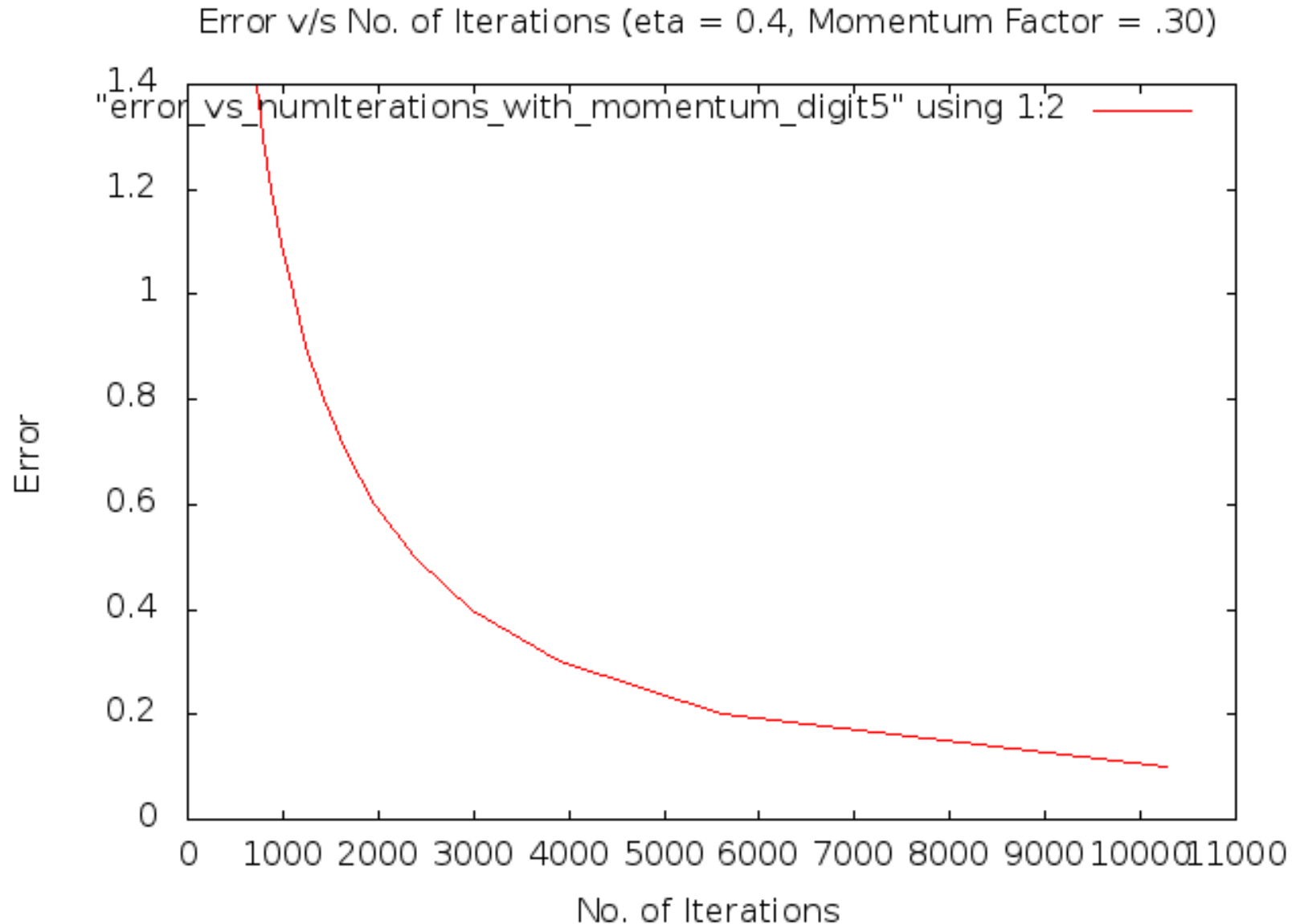
# Effect of Momentum Factor (Digit Recogniser)



# Effect of Momentum Factor (Digit Recogniser)

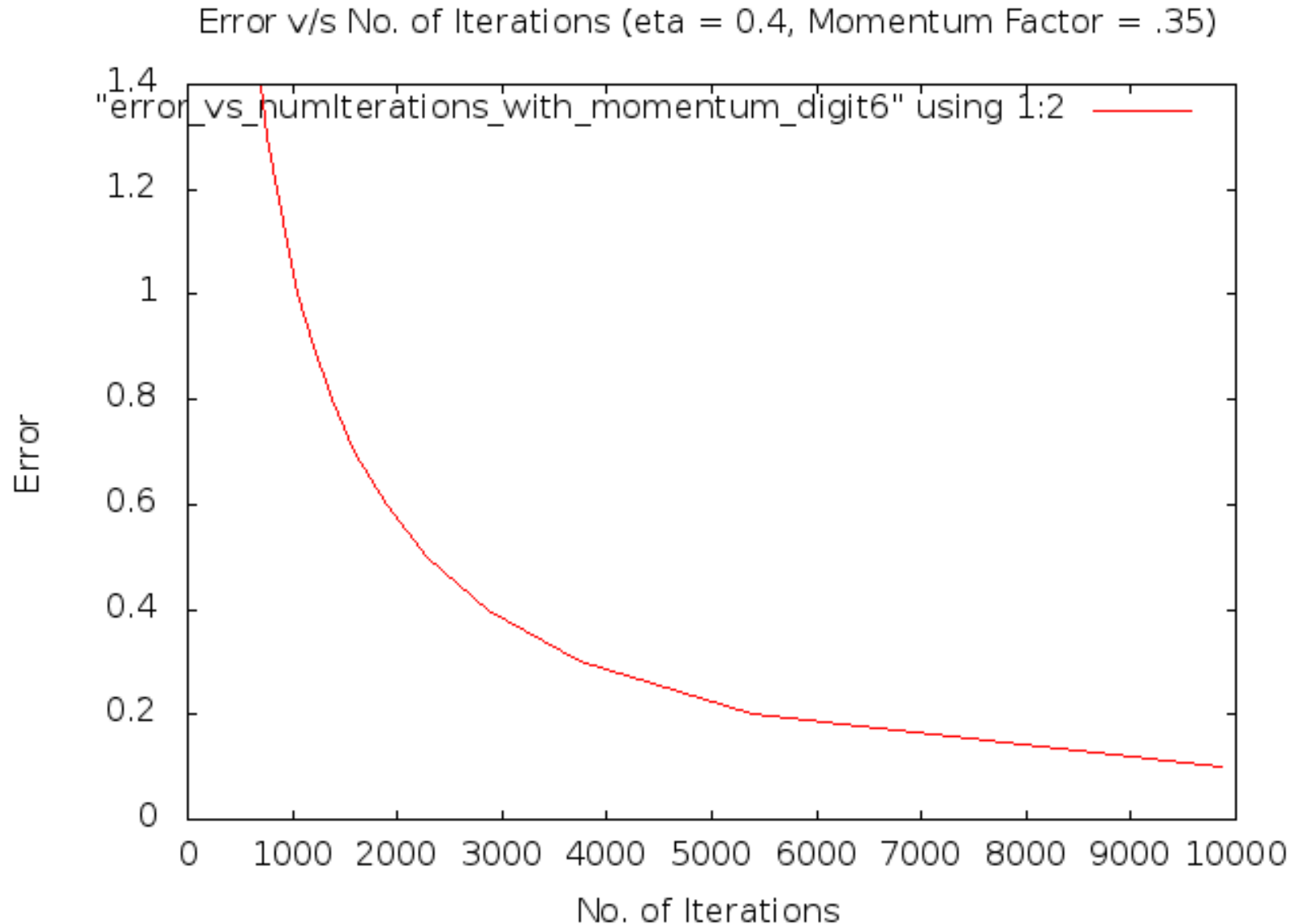


# Effect of Momentum Factor (Digit Recogniser)

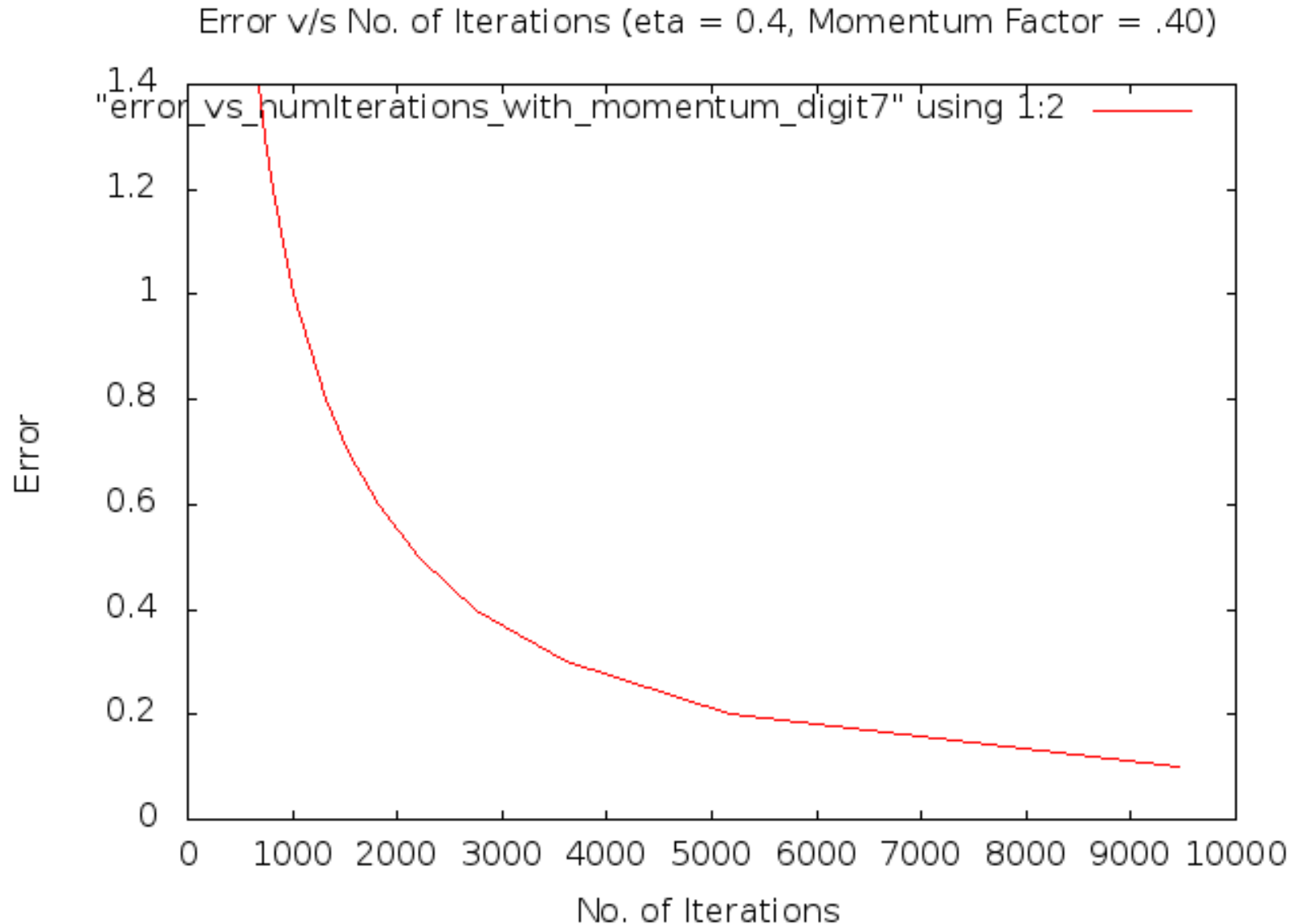




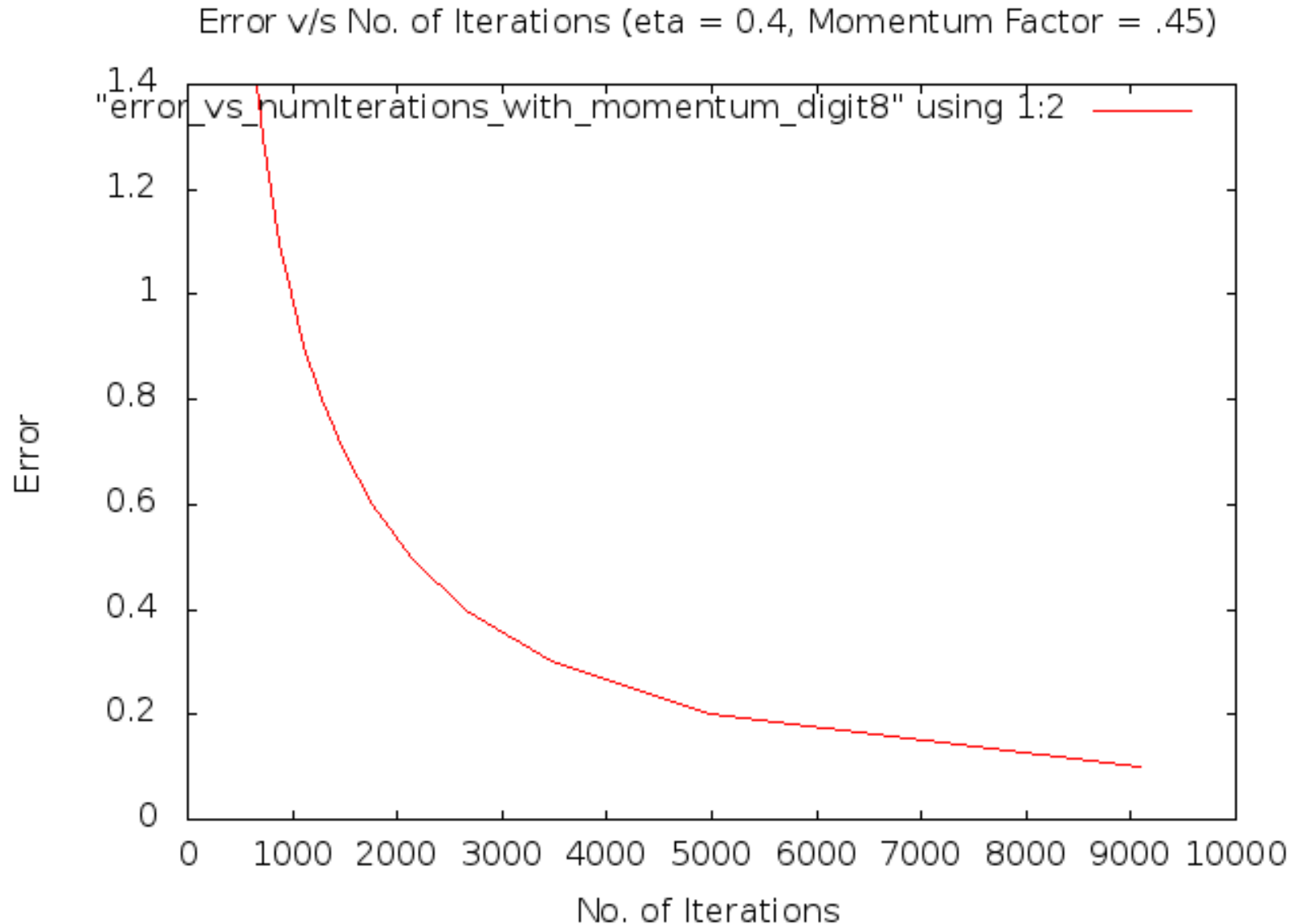
# Effect of Momentum Factor (Digit Recogniser)



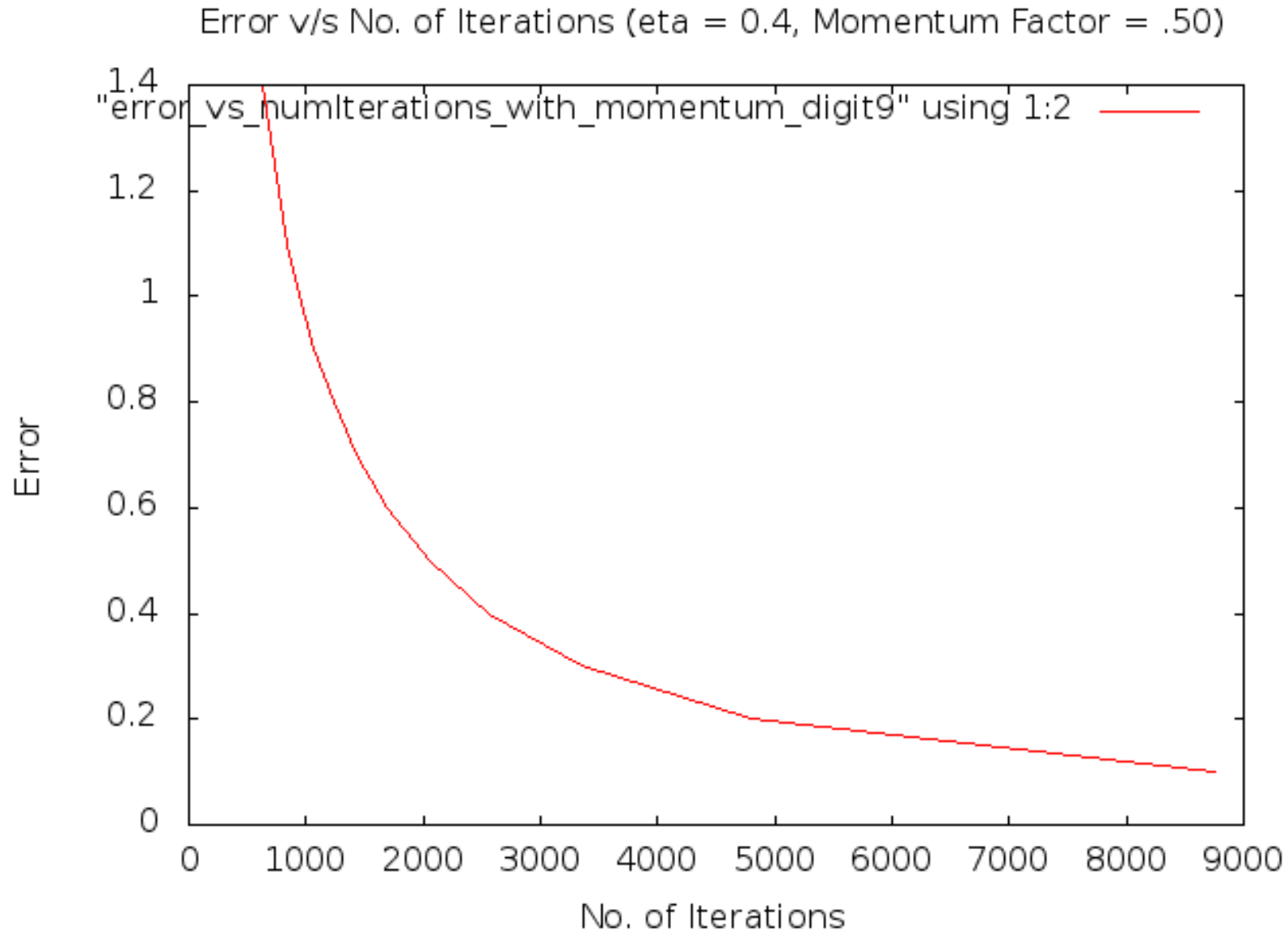
# Effect of Momentum Factor (Digit Recogniser)



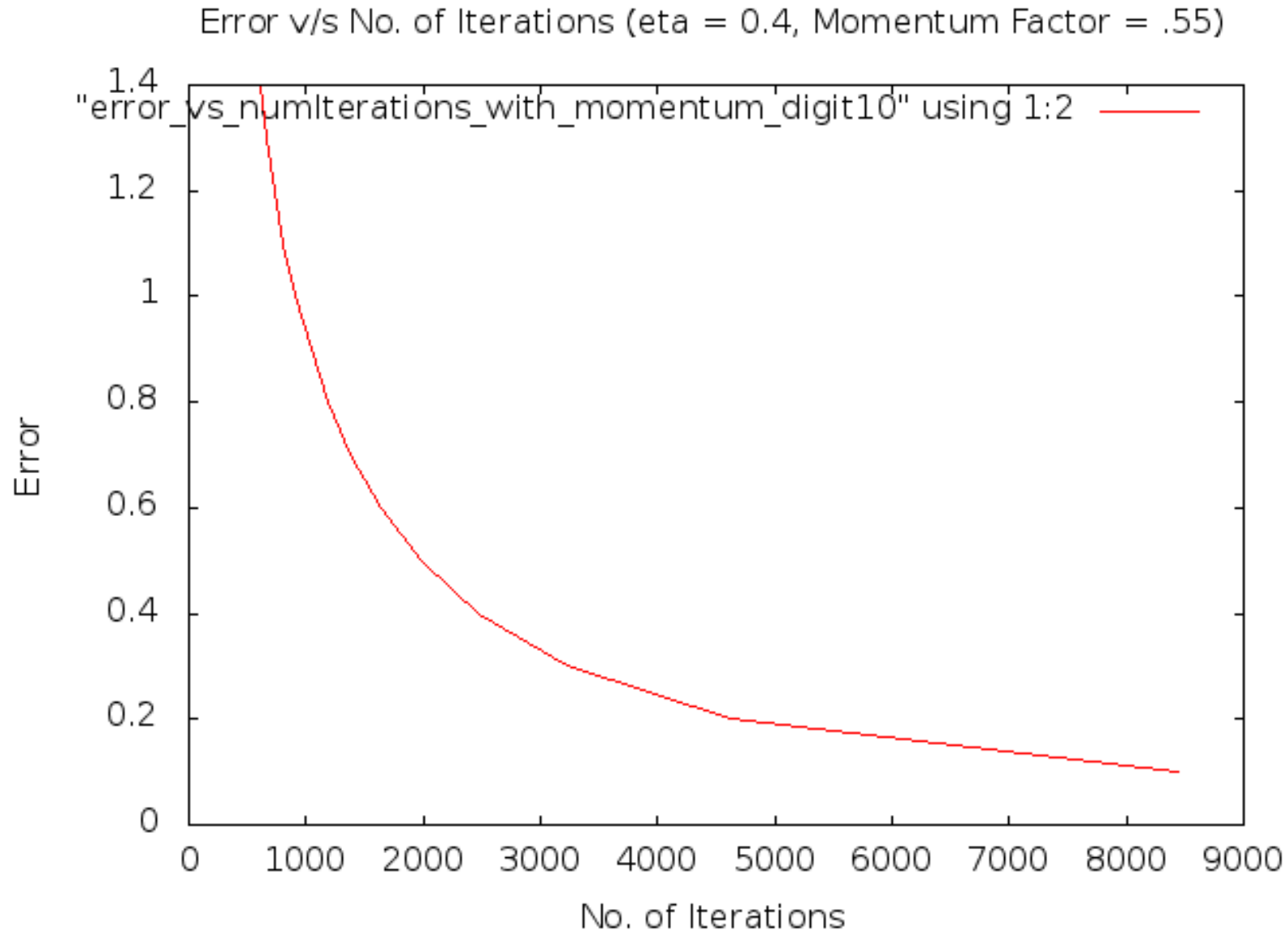
# Effect of Momentum Factor (Digit Recogniser)



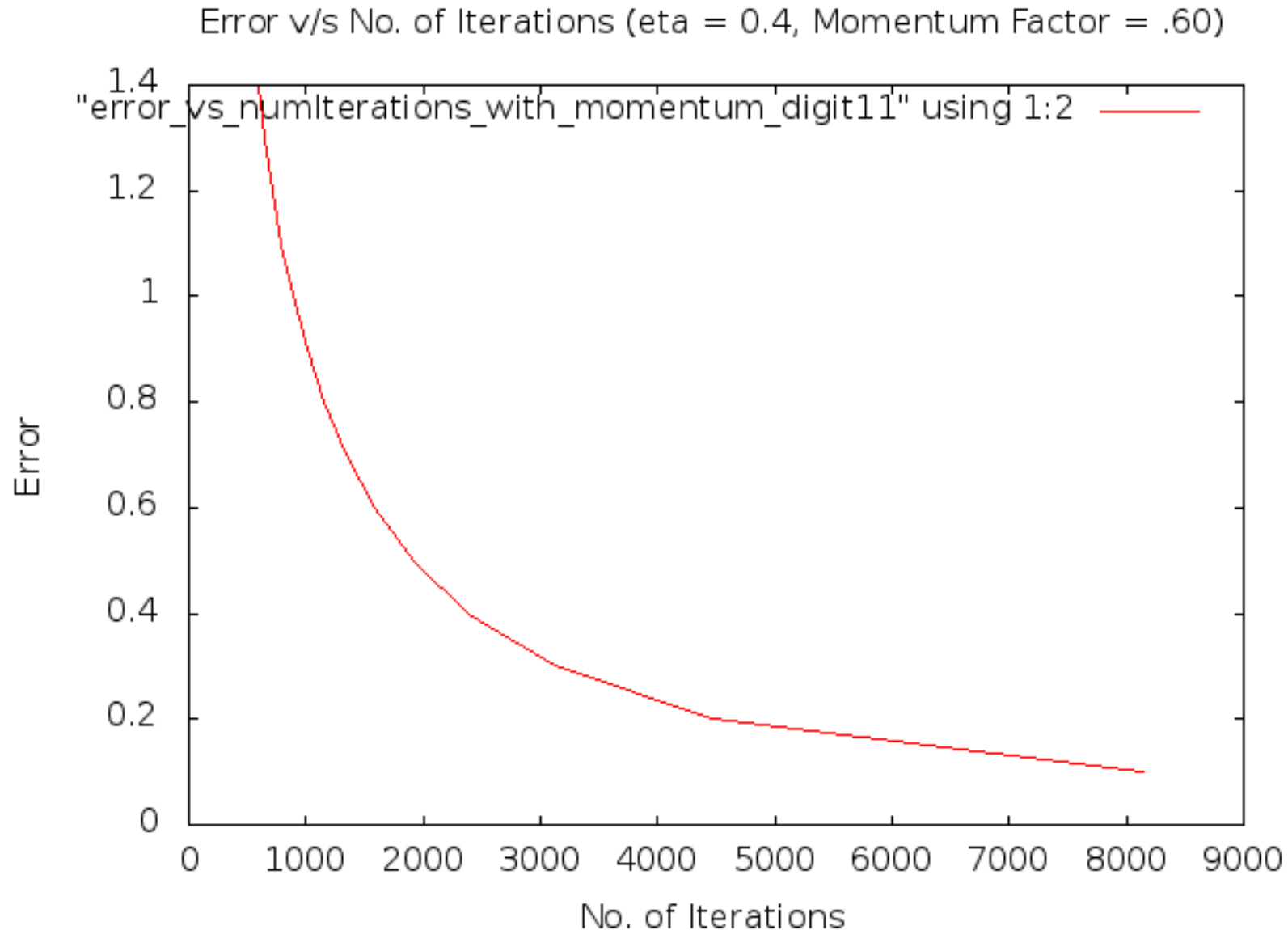
# Effect of Momentum Factor (Digit Recogniser)



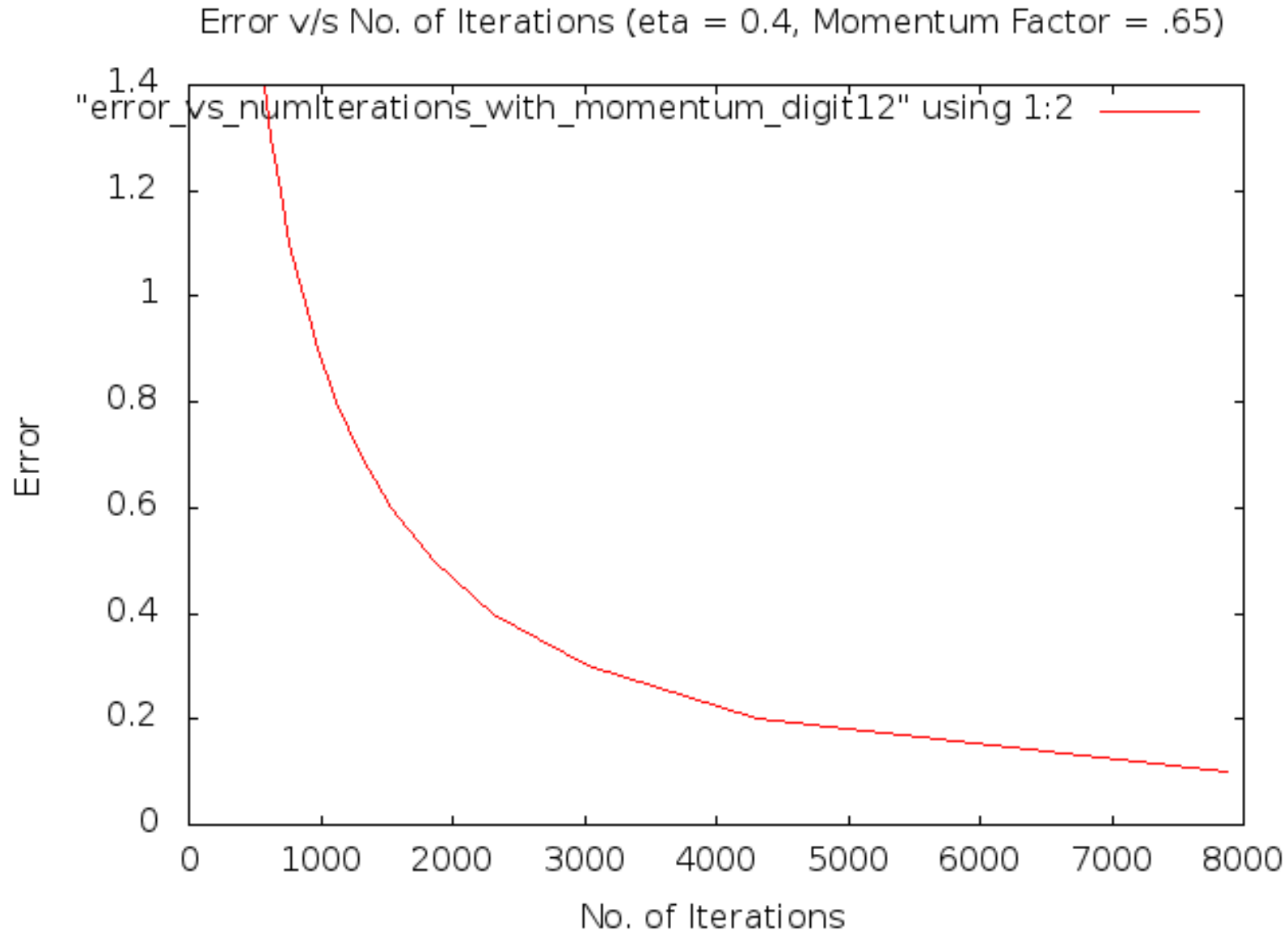
# Effect of Momentum Factor (Digit Recogniser)



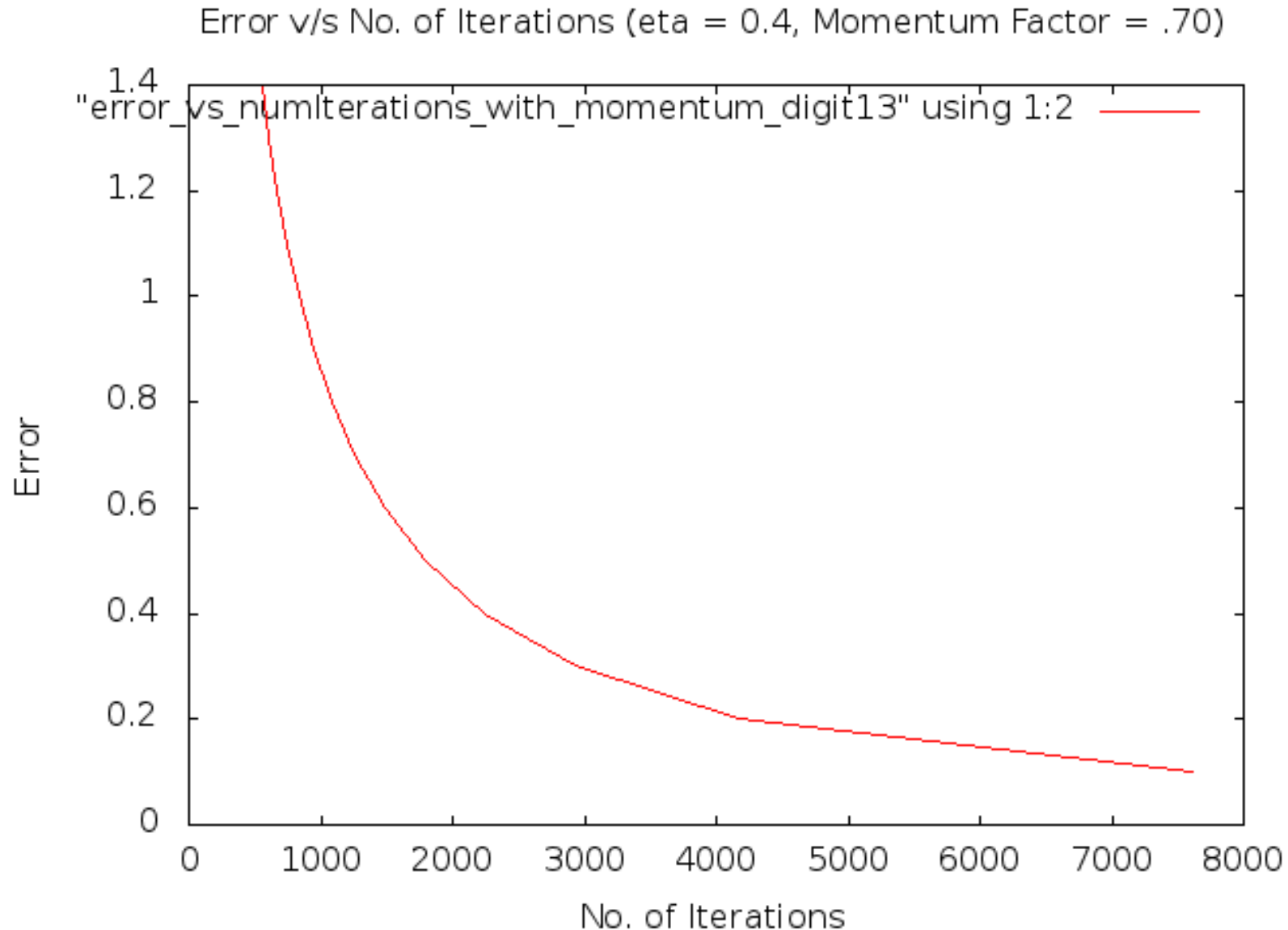
# Effect of Momentum Factor (Digit Recogniser)



# Effect of Momentum Factor (Digit Recogniser)

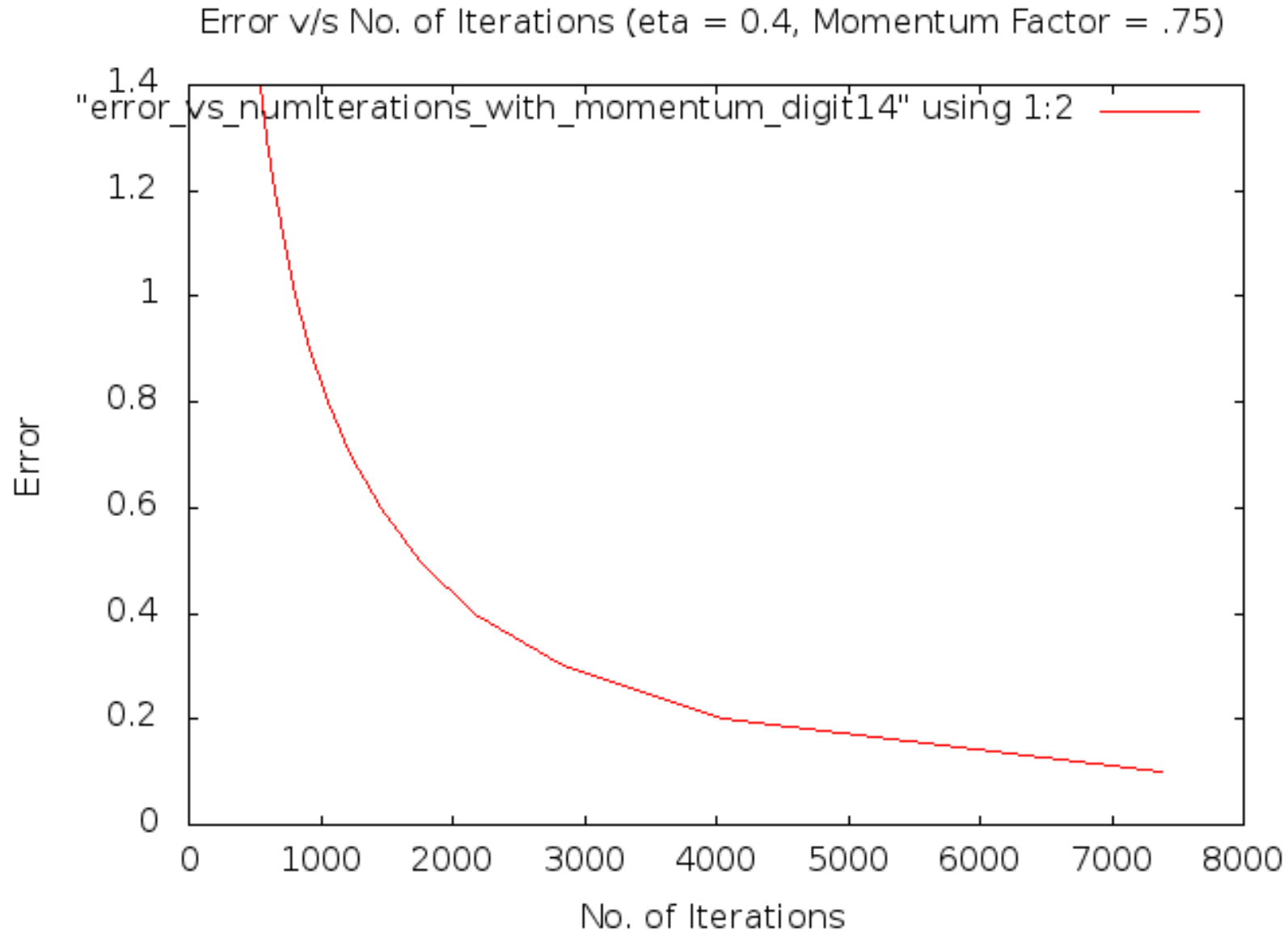


# Effect of Momentum Factor (Digit Recogniser)

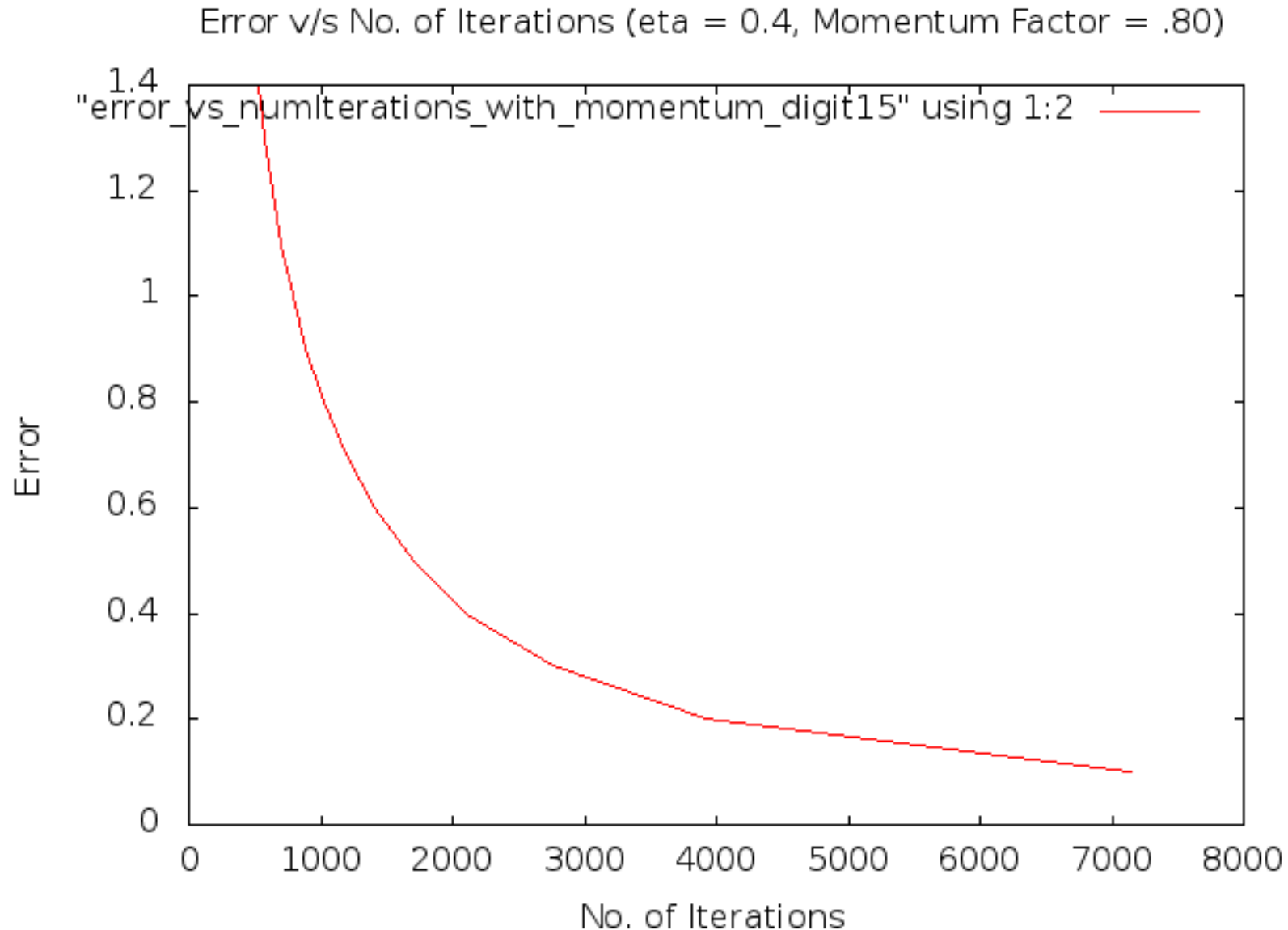




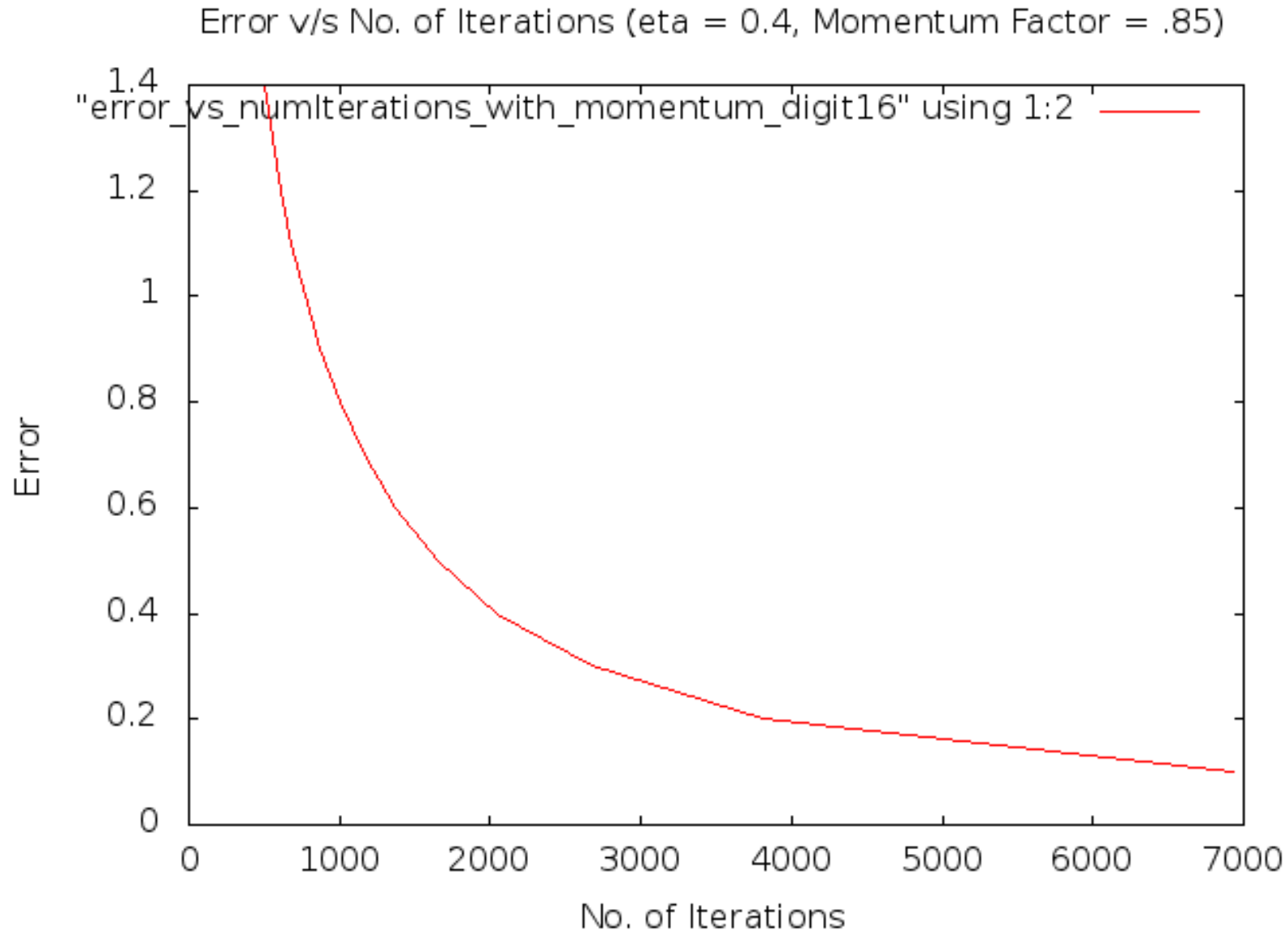
# Effect of Momentum Factor (Digit Recogniser)



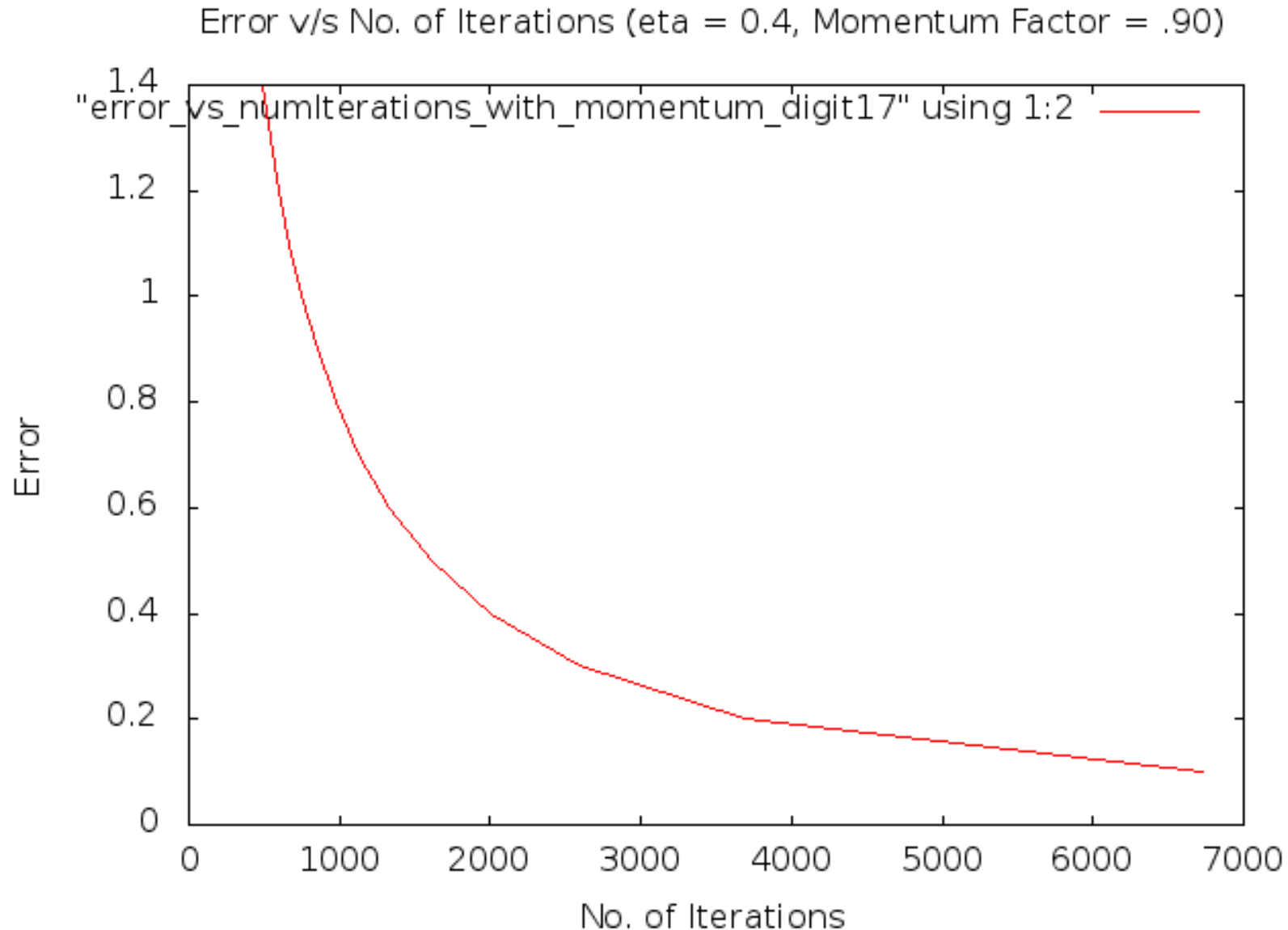
# Effect of Momentum Factor (Digit Recogniser)



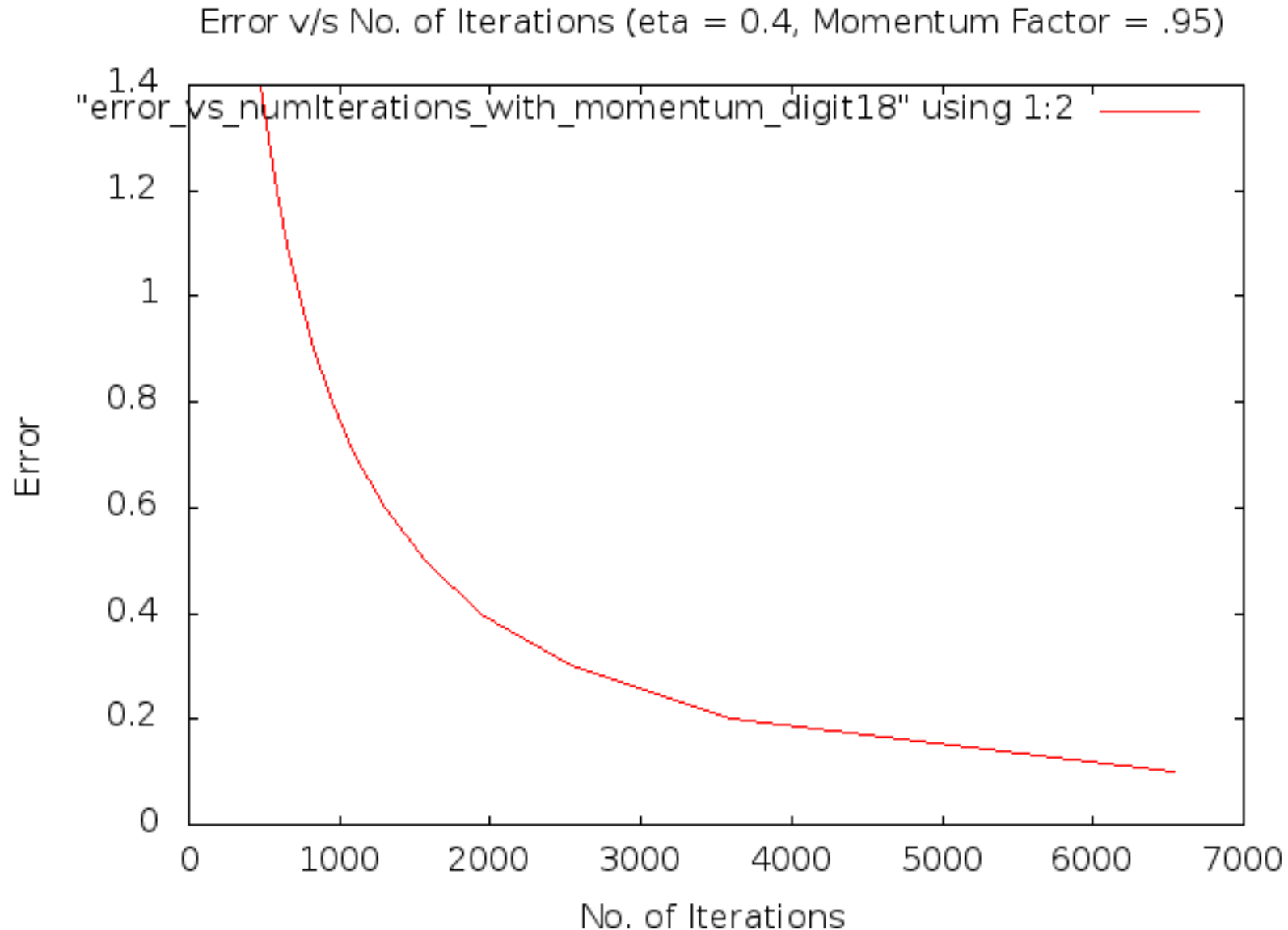
# Effect of Momentum Factor (Digit Recogniser)



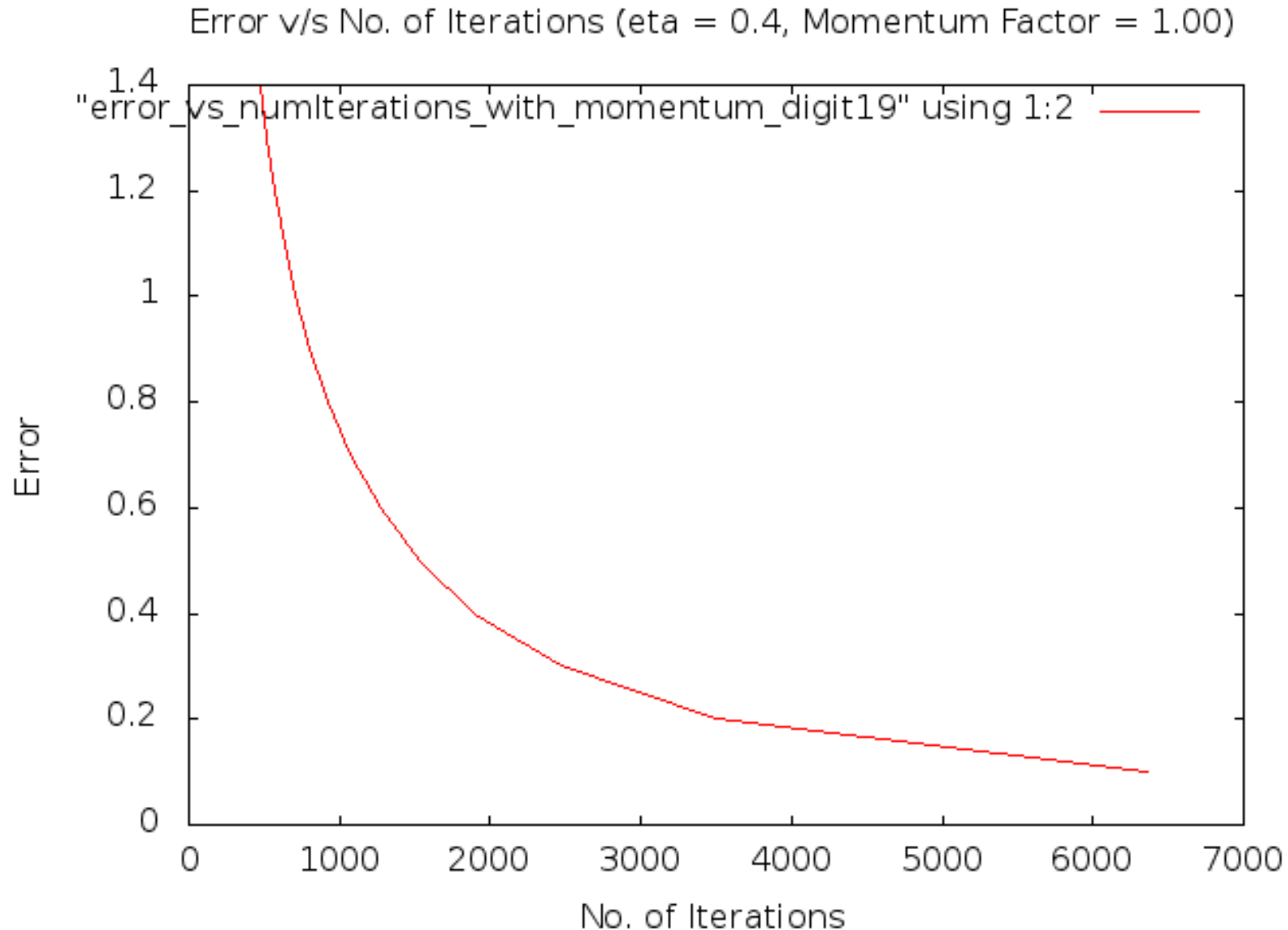
# Effect of Momentum Factor (Digit Recogniser)



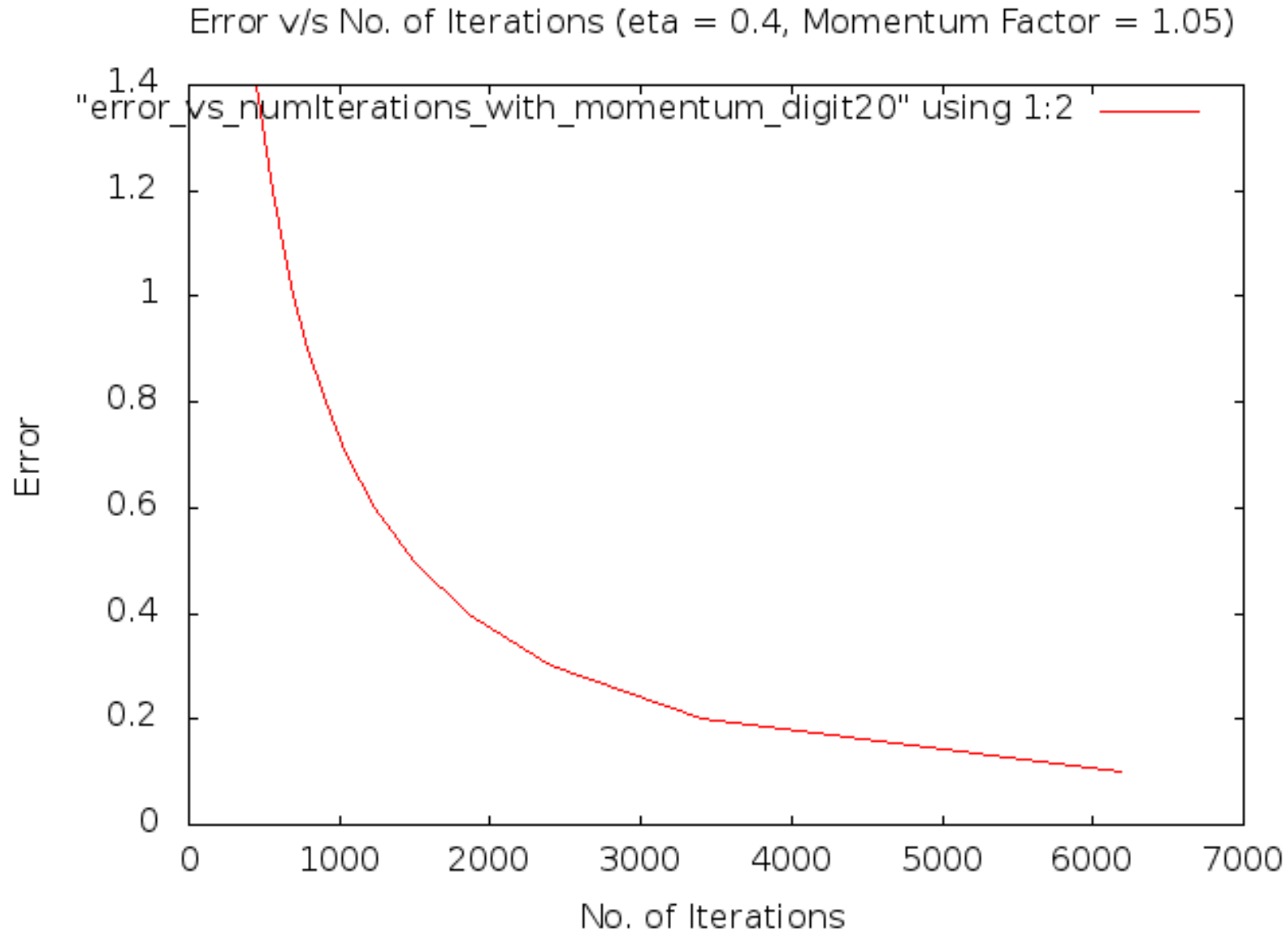
# Effect of Momentum Factor (Digit Recogniser)



# Effect of Momentum Factor (Digit Recogniser)

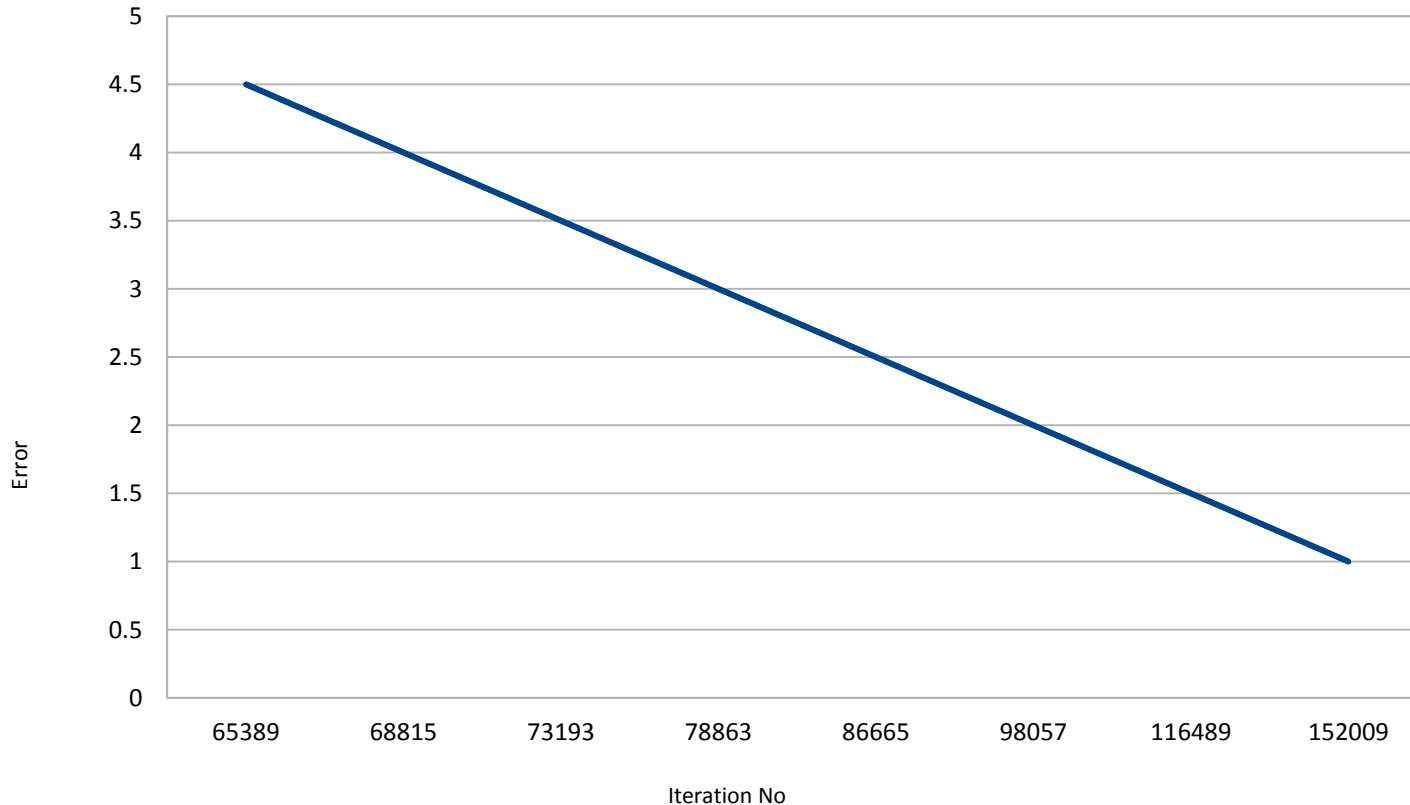


# Effect of Momentum Factor (Digit Recogniser)



# Effect of Momentum Factor (Parity)

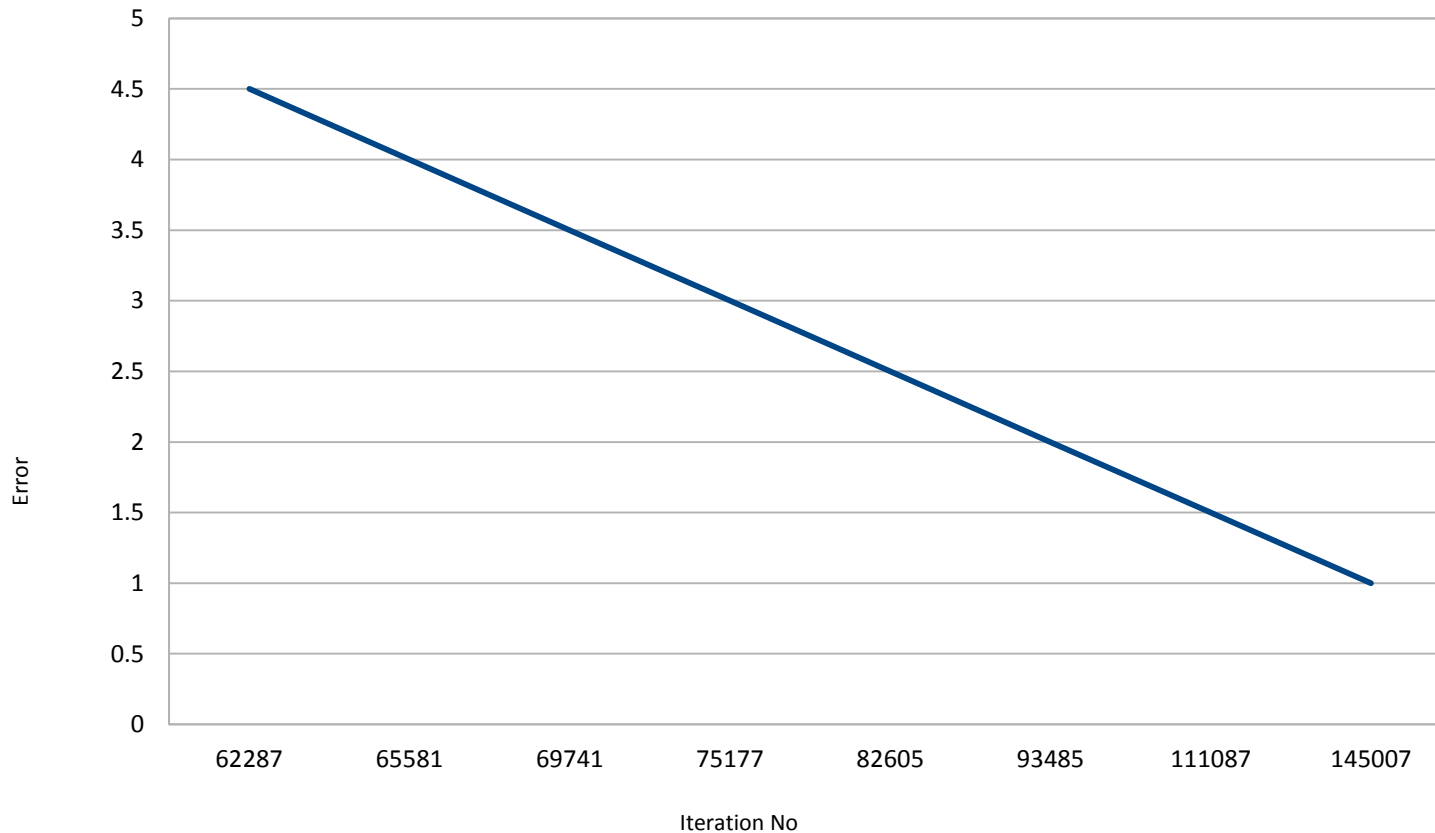
Eta=0.3 Momentum Factor=0.05





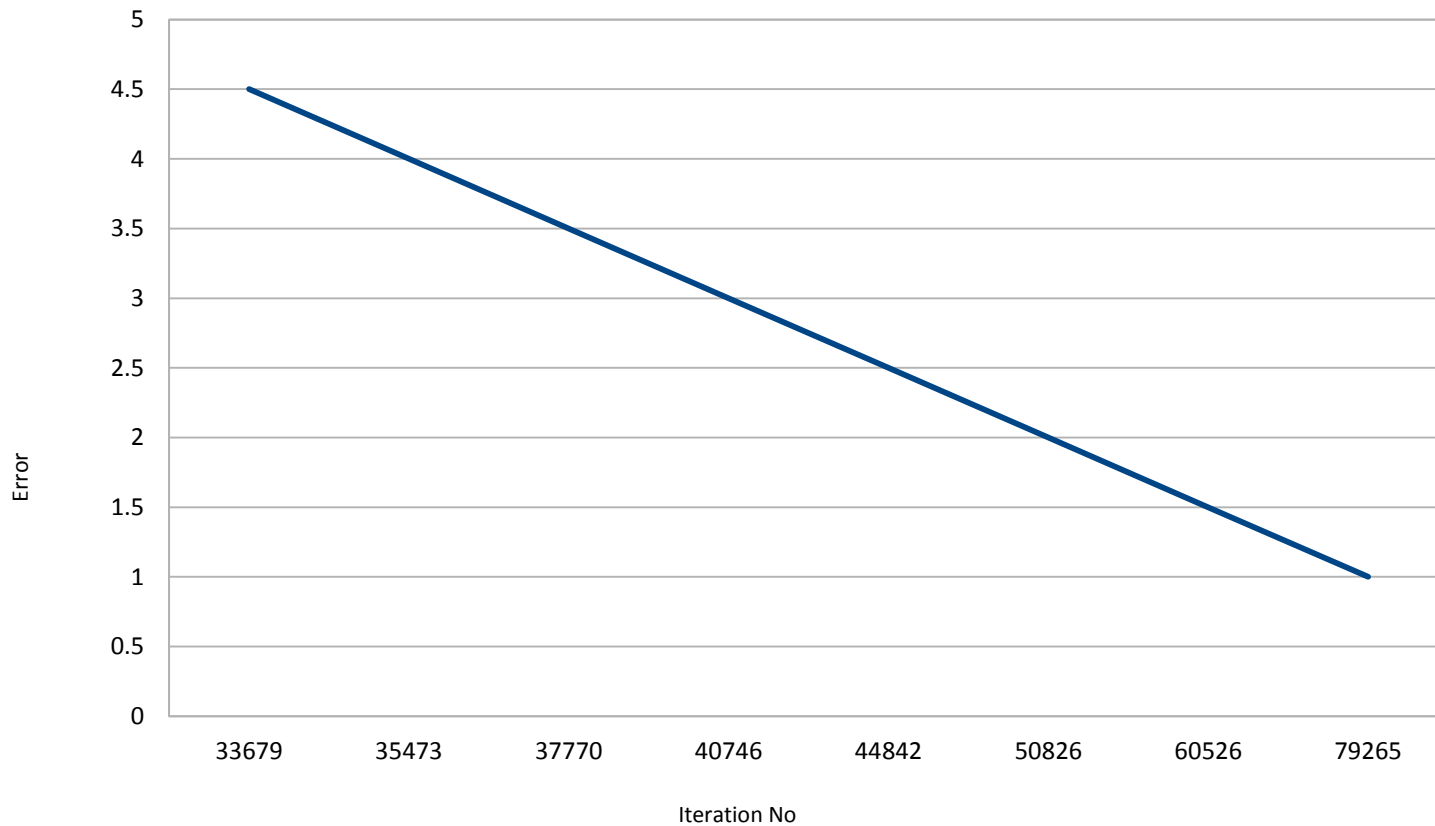
# Effect of Momentum Factor (Parity)

Eta=0.3 Momentum Factor=0.1



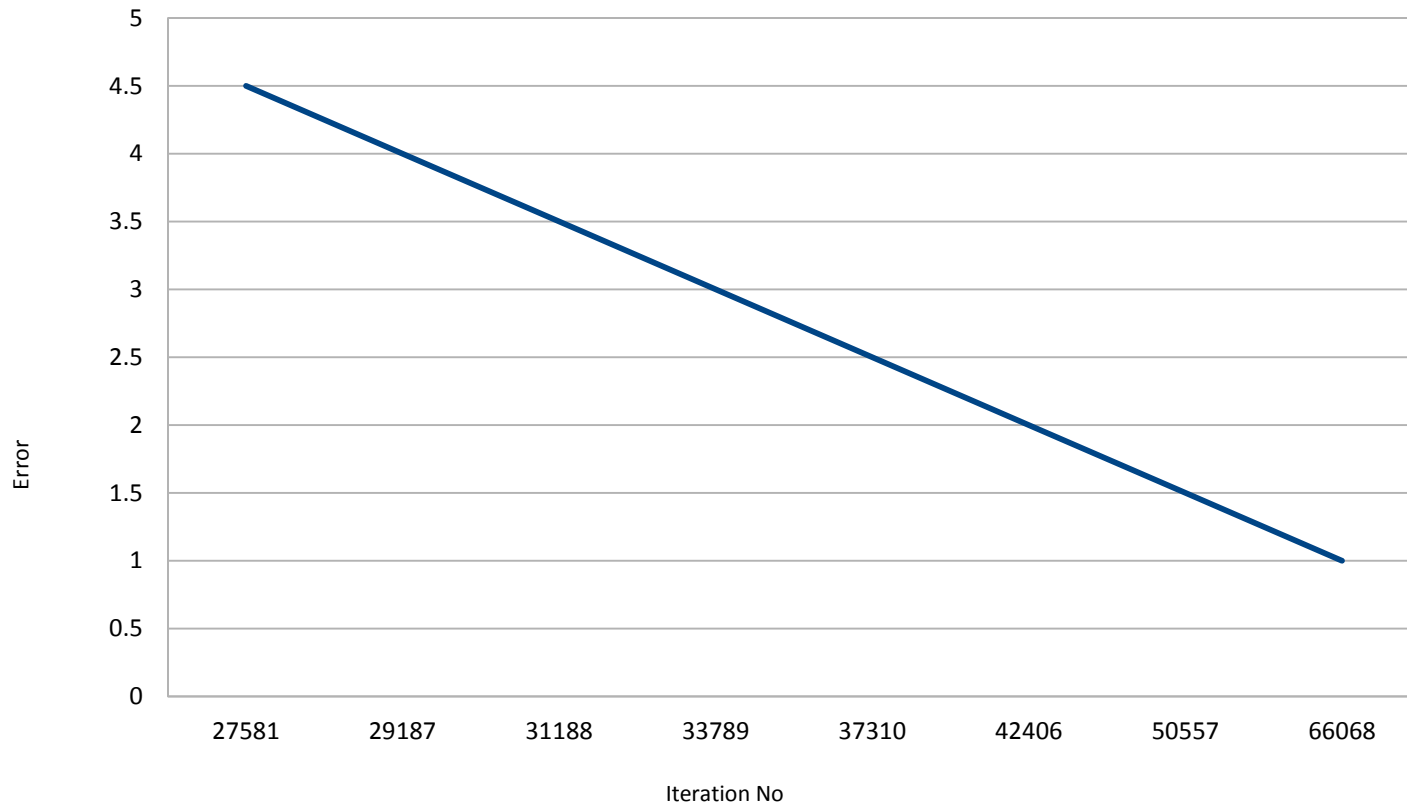
# Effect of Momentum Factor (Parity)

Eta=0.3 Momentum Factor=1



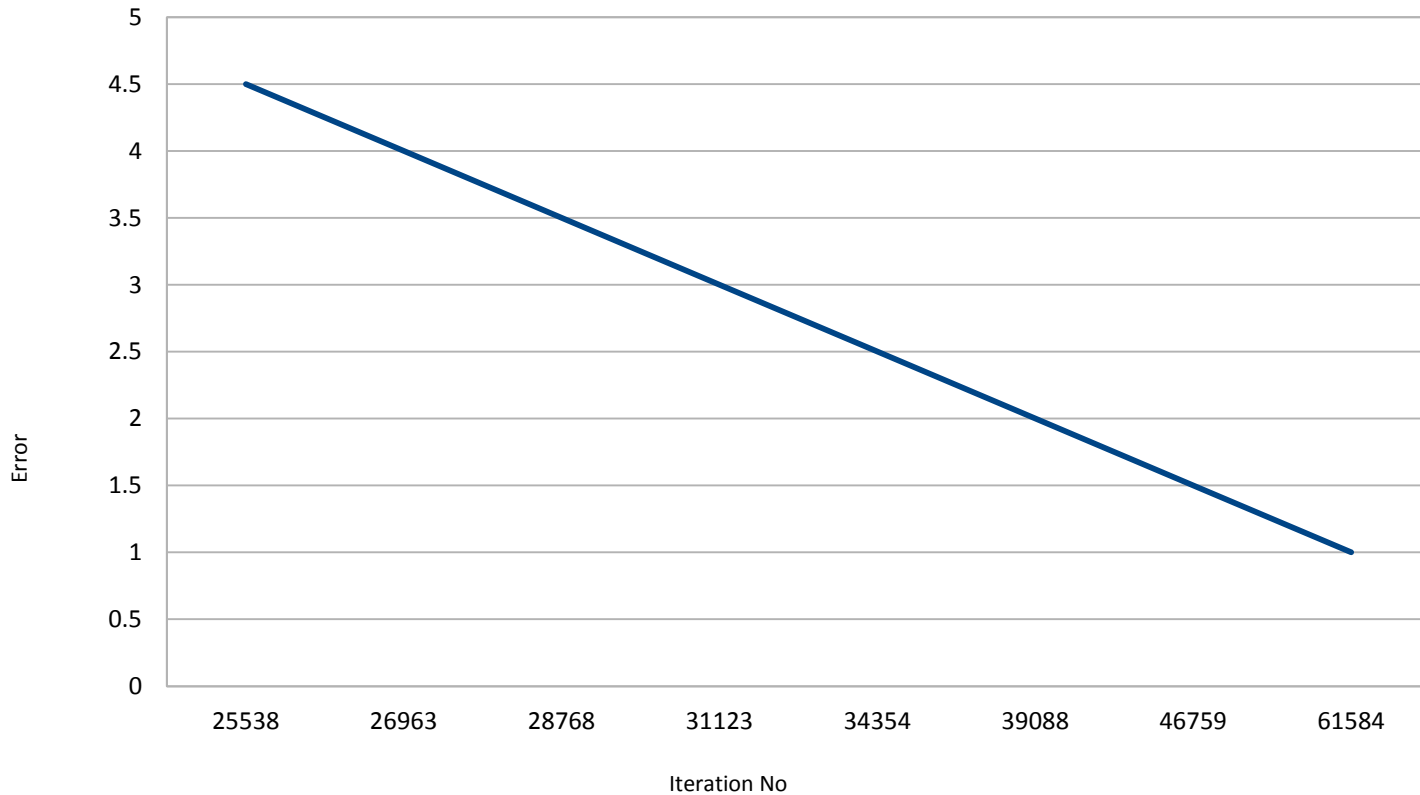
# Effect of Momentum Factor(Palindrome)

Eta = 0.3 Momentum Factor = 0.05



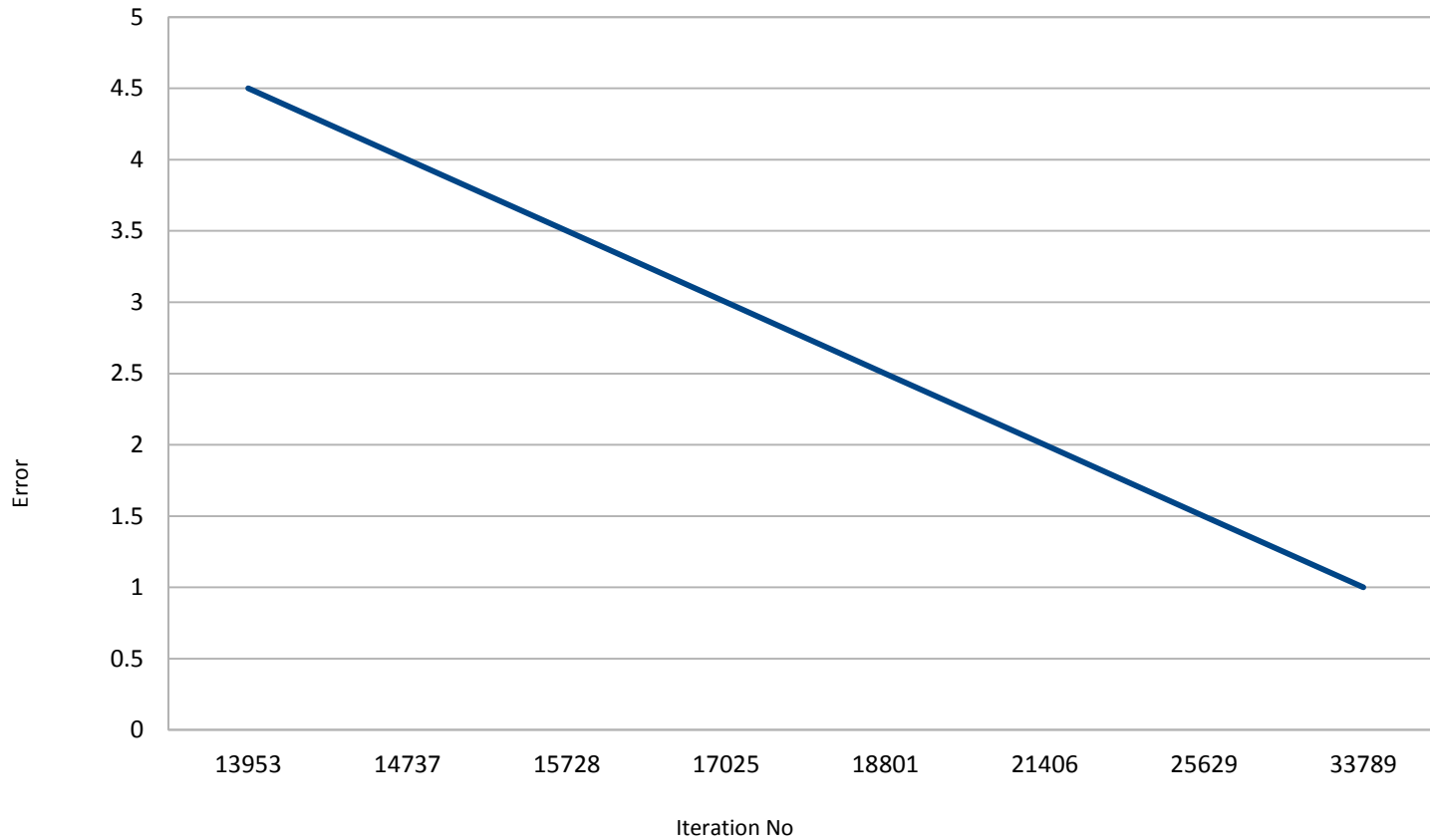
# Effect of Momentum Factor(Palindrome)

Eta = 0.3 Momentum Factor = 0.1



# Effect of Momentum Factor(Palindrome)

Eta = 0.3 Momentum Factor = 1



# Effect of Momentum Factor

- We observed the effect of increasing the momentum too from close to 0 (0.001) to close to 1(1.05) in steps of 0.05
- In general, as the momentum factor increased the no of iterations required to converge to a solution also decreased.
- We also observed that as the momentum factor increases, iteration-to-iteration variation in error was greatly increased.

# Functionality of Hidden Layer Neurons

## 2-Input XOR

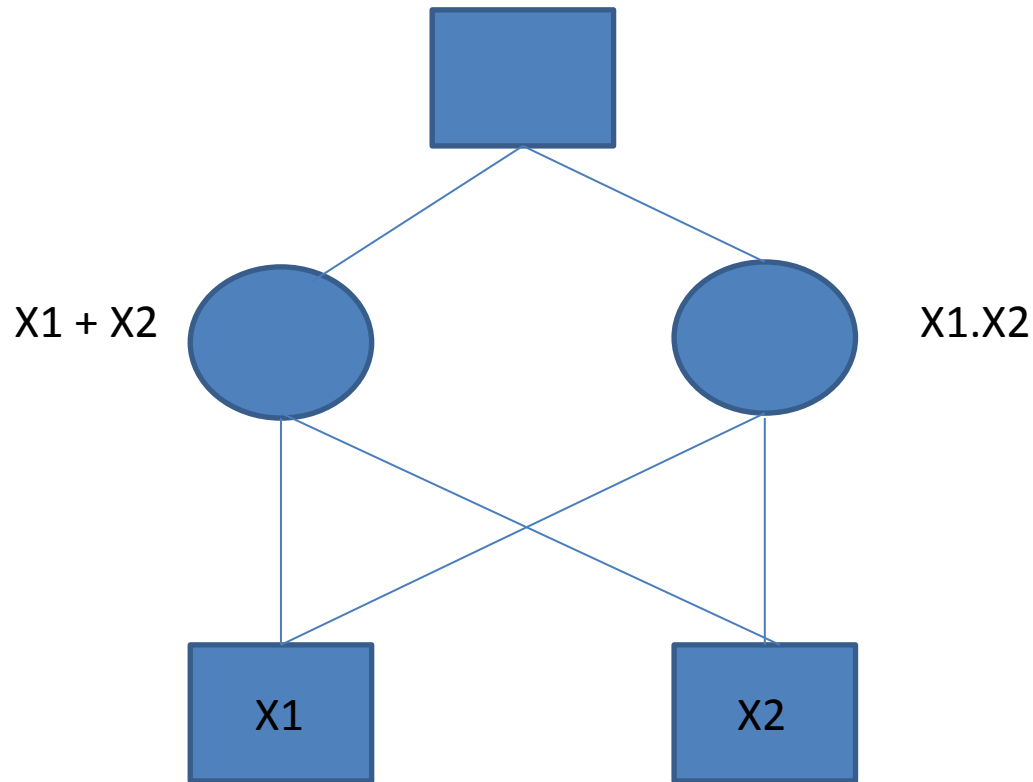
Input1	Input2	HiddenLayerOp1	HiddenLayerOp2	Output
0	0	0	0	0
0	1	1	0	1
1	0	1	0	1
1	1	1	1	0

- $\text{HiddenLayerOp1} = \text{Input1} \text{ OR } \text{Input2}$
- $\text{HiddenLayerOp2} = \text{Input1} \text{ AND } \text{Input2}$
- $\text{Output} = \text{HiddenLayerOp1} \text{ AND } (\text{NOT} (\text{HiddenLayerOp2}))$

# Functionality of Hidden Layer Neurons

## 2-Input XOR

$$(X1 + X2) \cdot \overline{(X1 \cdot X2)} = X1 \cdot \overline{X2} + X2 \cdot \overline{X1}$$





# Functionality of Hidden Layer Neurons

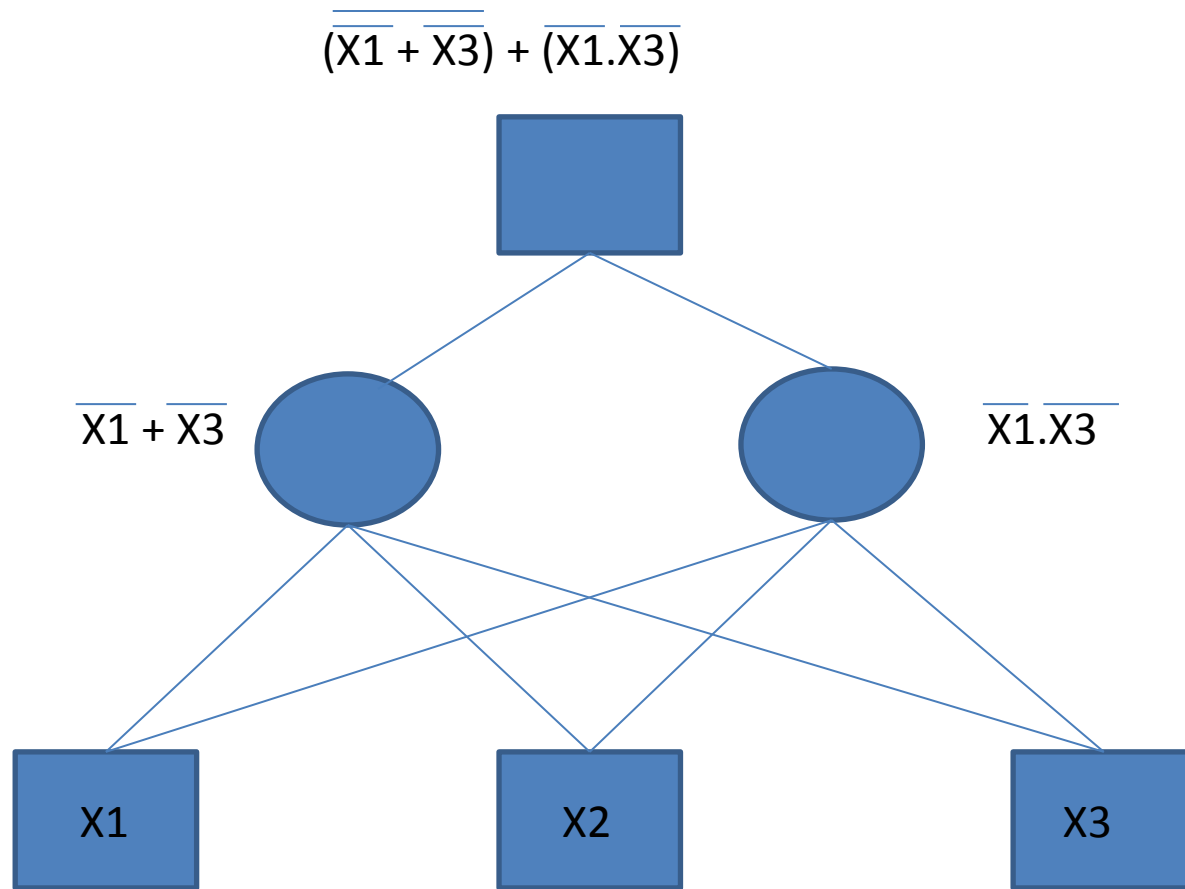
## 3-Input Palindrome

Input1	Input2	Input3	HiddenLayerOp1	HiddenLayerOp2	Output
0	0	0	1	1	1
0	0	1	1	0	0
0	1	0	1	1	1
0	1	1	1	0	0
1	0	0	1	0	0
1	0	1	0	0	1
1	1	0	1	0	0
1	1	1	0	0	1

- HiddenLayerOp1: NOT(Input1) OR NOT(Input3)
- HiddenLayerOp2 : NOT(Input1) AND NOT(Input3)
- Output = NOT(HiddenLayerOp1) OR HiddenLayerOp1

# Functionality of Hidden Layer Neurons

## 3-Input Palindrome



For Palindrome, the  $X2$  should not matter. Interestingly,  $X2$  is not even taken into consideration in hidden layer.

# Functionality of Hidden Layer Neurons

## 5-Input Palindrome

Input 1	Input 2	Input 3	Input 4	Input 5	HiddenLayerO p1	HiddenLayerO p2	Output
0	0	0	0	0	0	0	1
0	0	0	0	1	1	0	0
0	0	0	1	0	1	0	0
0	0	0	1	1	1	0	0
0	0	1	0	0	0	0	1
0	0	1	0	1	1	0	0
0	0	1	1	0	1	0	0
0	0	1	1	1	1	0	0

Error Threshold = 5%

Initial Weights = Random between -1 & 1

Learning Rate = 0.34

# Functionality of Hidden Layer Neurons

## 5-Input Palindrome

Input 1	Input 2	Input 3	Input 4	Input 5	HiddenLayerO p1	HiddenLayerO p2	Output
0	1	0	0	0	0	1	1
0	1	0	0	1	0	1	0
0	1	0	1	0	0	0	0
0	1	0	1	1	1	0	0
0	1	1	0	0	0	1	1
0	1	1	0	1	0	1	0
0	1	1	1	0	0	0	0
0	1	1	1	1	1	0	0

Error Threshold = 5%

Initial Weights = Random between -1 & 1

Learning Rate = 0.34

# Functionality of Hidden Layer Neurons

## 5-Input Palindrome

Input 1	Input 2	Input 3	Input 4	Input 5	HiddenLayerO p1	HiddenLayerO p2	Output
1	0	0	0	0	0	1	0
1	0	0	0	1	0	0	1
1	0	0	1	0	1	0	0
1	0	0	1	1	1	0	0
1	0	1	0	0	0	1	0
1	0	1	0	1	0	0	1
1	0	1	1	0	1	1	0
1	0	1	1	1	1	0	0

Error Threshold = 5%

Initial Weights = Random between -1 & 1

Learning Rate = 0.34

# Functionality of Hidden Layer Neurons

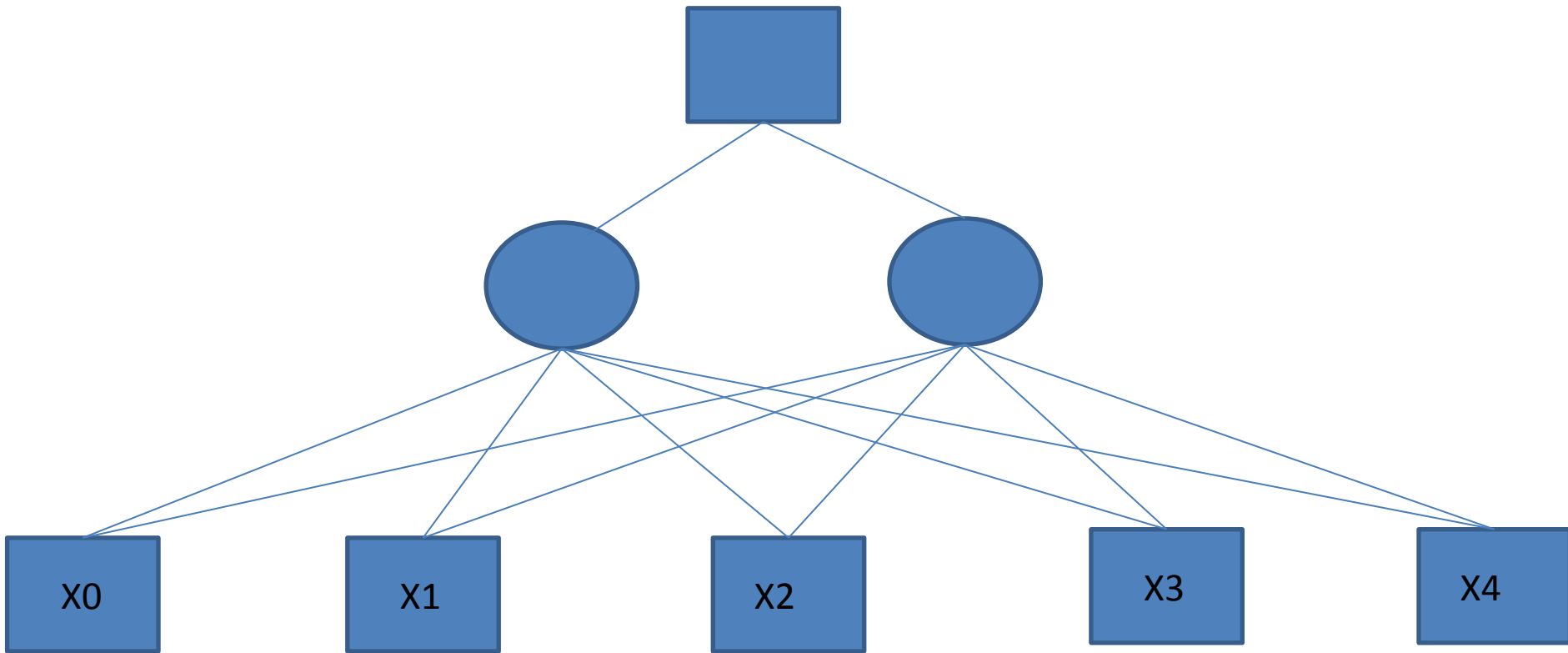
## 5-Input Palindrome

Input 1	Input 2	Input 3	Input 4	Input 5	HiddenLayerO p1	HiddenLayerO p2	Output
1	1	0	0	0	0	1	0
1	1	0	0	1	0	1	0
1	1	0	1	0	0	1	0
1	1	0	1	1	0	0	1
1	1	1	0	0	0	1	0
1	1	1	0	1	0	1	0
1	1	1	1	0	0	1	0
1	1	1	1	1	0	0	1

After Solving through Karnaugh Map, we get  
Output = (H1 NOR H2)

# Functionality of Hidden Layer Neurons

## 5-Input Palindrome



For Palindrome, the  $x_2$  should not matter. Interestingly,  $x_2$  is not even taken into consideration in hidden layer.