Module Code: CS2NC19
Assignment report Title: Neurocomputation Conference Paper
Student Number: 26011251
Date (when the work completed): 26/02/20
Actual hrs spent for the assignment: 36

1. Abstract

The three experiments using RGB colour code data sets have allowed conclusions to be drawn on the effects a changing learning rate, momentum, and the number of hidden neurons has on the final SSE value and the number of epochs required to reach this value. It has been shown an increasing learning rate will result in fewer epochs required however between learning rate values of 0.5 and 0.6 the number of epochs increases, suggesting the best course of action is not just to use the high learning rate possible. It has also been shown momentum has not had an effect on learning the colour codes as the second experiment produced the same results as the first. Lastly it has been shown there is no correlation between the number of hidden neurons and the number of epochs required to reach the final SSE value.

2. Introduction

A neural network is a series of algorithms that endeavours to recognise underlying relationships in a set of data through a process that mimics the way the human brain operates [1]. Multi-layer-perceptrons input data which is then manipulated by set weights in order to produce a target output, these weights are then adjusted to reach the target output and reduce error. These networks can be used to classify, associate, or predict, such as in weather systems.

Through three practical sessions a neural network has been programmed in Java. Java is an object-oriented language, allowing for the creation of unique classes with their own attributes. These classes can be inherited by subclasses, reducing the amount of code required and allowing for complicated systems such as the multi-layer-perceptron.

RGB colour codes have been chosen as the data to be used in the experiments. Each colour code has three inputs, corresponding to each colour value between 0-255. Each code is then assigned a target to meet based upon their colour (IE. all reds have a target of 0).

The experiments will test to measure the final SSE value and the number of epochs required to reach this value. This will show the effects, if any, of each of the three chosen parameters. It is expected that a high learning rate will allow for quicker training of the network [2]. It is also expected that there will be an optimal number of hidden neurons, after which the hidden layer will be "over-fitted" and the network will not be able to interpret the data properly [3].

3. User Application

To explore the effect changing the Learning Rate, Momentum, and number of Hidden Neurons has on the Neural Network's ability to learn, suitable data must be gathered. This data has to be continuous, and a small change in the input should result in a small change in the output; this allows the neural network to generalise and interpolate.

There also has to be enough data to allow for three unique data sets: the validation set, the testing set, and the training set.

The validation set provides an unbiased evaluation of a model fit on the training dataset while tuning the model's hyper parameters eg. Number of Hidden Neurons [4]. The testing set provides an unbiased evaluation of a final model fit on the training dataset [4]. The training set is a set of examples used to fit the parameters of the model [5].

For ease of constructing these three sets, RGB colour codes were chosen for the data set as each colour has three attributes, each of these attributes is continuous, and there are 16777216 different combinations of colours. Consequently it is easy to gather data for each of the sets.

The RGB colour codes were acquired from htmlcolorcodes.com [6]. The neural network would not allow for all 140 named colours to be used a large amount of data would cause Index errors in the code, as such 44 different RGB values are used in each data set (9 red, 5 pink, 11 yellow, and 19 purple) . The validation set uses the RGB values from the website. Both the training and testing sets use the values of the validation set +-10, alternating for each value.

There are three input neurons, each input uses either the R, the G, or the B value.Colours of the same type (IE. light coral and salmon are both red) share the same target, allowing the neural network to use multiple RGB values to learn one type of colour. It was originally planned to test the effect of grouping colours together and then with unique targets however setting each colours' target to a unique value caused Index errors in the code.

4. Results of the Application

Three experiments were performed to determine the effect of the learning rate, momentum, and the number of hidden neurons. Each experiment tested the number of training epochs against the Sum of Squared Errors (SSE). The neural network was set to learn the colours using the data sets provided.

The final SSE was determined by calculating the mean average of the last 10 SSEs and comparing to the current SSE, if the current SSE is greater than the previous average then learning will stop as the error should not increase.

Each parameter increased by set amounts ten times. Graphs of Epochs vs SSE were then constructed for each experiment.

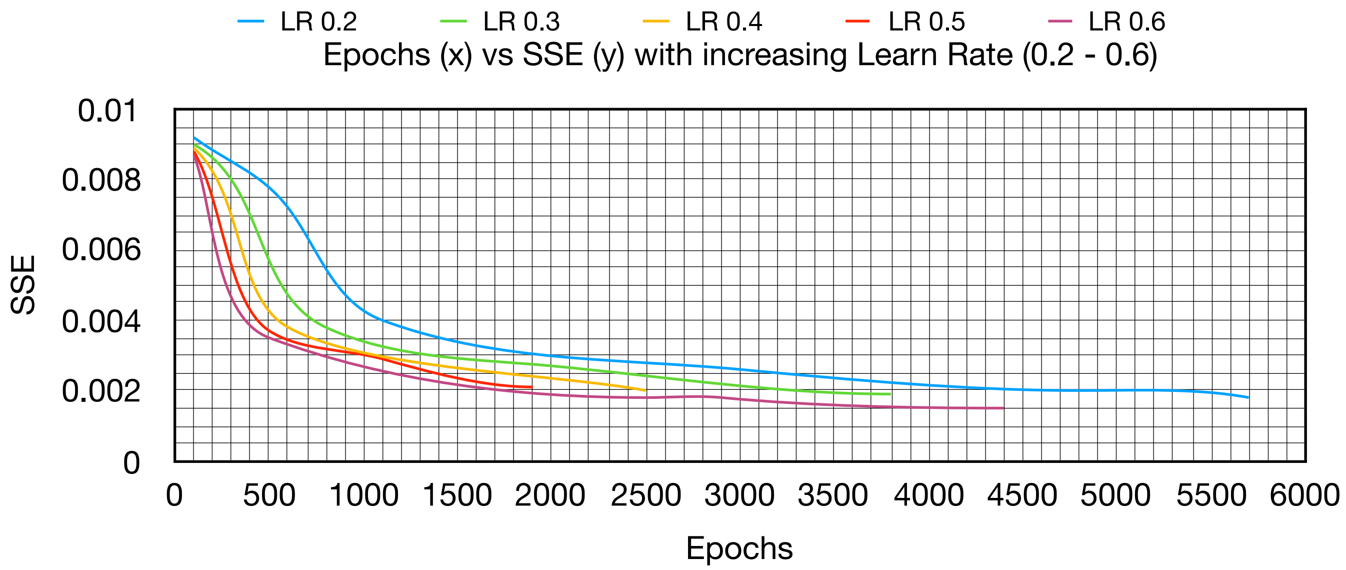
4.1 Experiment One: Learning Rate

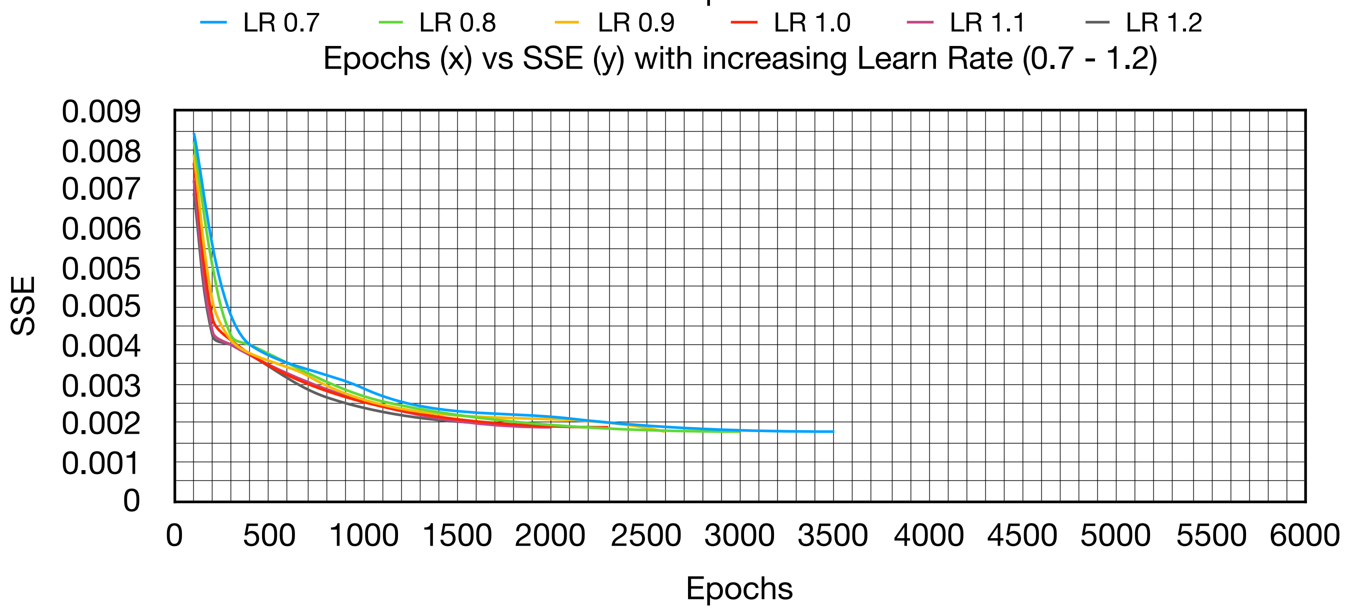
The first experiment considered an increasing learning rate. The initial learning rate was set to 0.2, increasing by 0.1 with each successive test to a maximum value of 1.2. For clarity between lines on the graphs, the results have been split into two graphs.

The first graph shows an overall trend, as the learning rate increases, the number of epochs decreases. The final SSE is found within a range of 0.0015 to 0.0020. The number of epochs decreases from 5760 to 1960 at a learning rate of 0.5. The result for a learning rate of 0.6 has a greater number of epochs at 4490 but has the lowest final SSE of 0.0015.

The second graph also follow the same trend as the first graph, as the learning rate increases, the number of epochs decreases. The final SSE is found within a range of 0.0016 to 0.0018, this is a lower range than the first graph. The number of epochs decreases from 3560 to 1720 at a learning rate of 1.2, this is the lowest number of epochs however it does not have the lowest SSE.

The overall range of final SSE values is 0.0015 to 0.0020. The gradient of the second graph is steeper than the first graph, showing a greater change in SSE values as the learning rate increases. Furthermore the greatest change in SSE is found within the first 500 epochs. Learn rate is not directly proportional to the number of epochs.





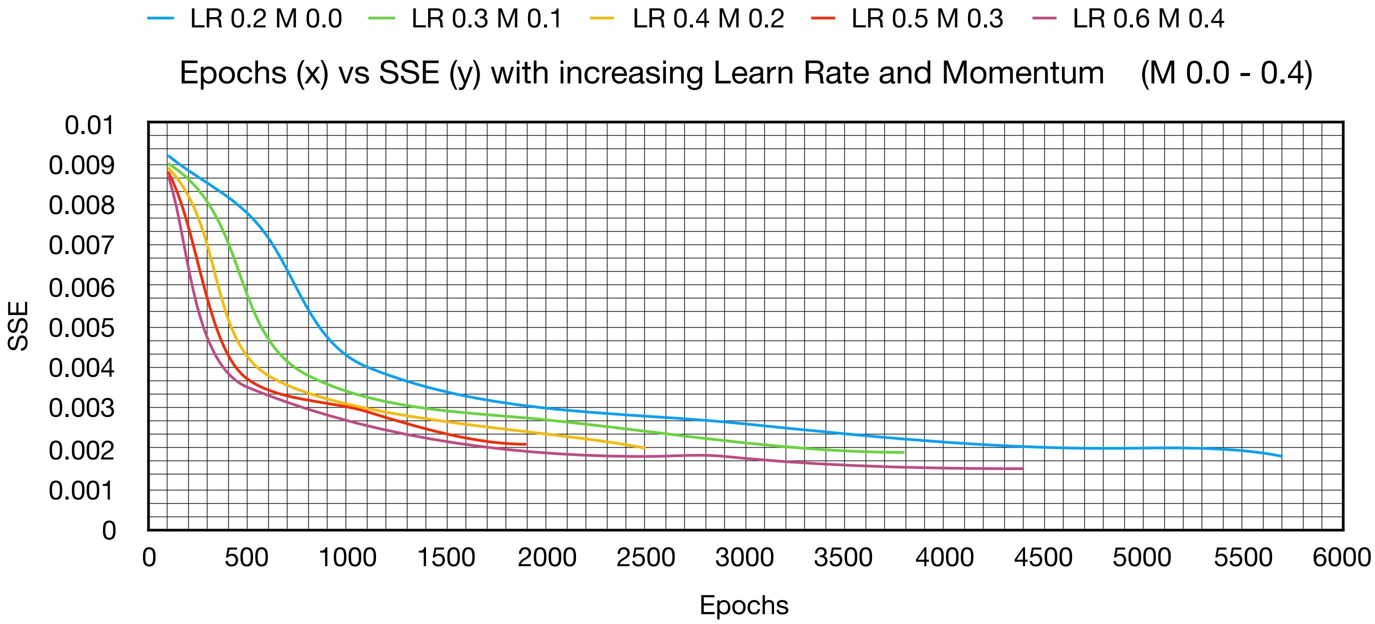
4.2 Experiment Two - Momentum

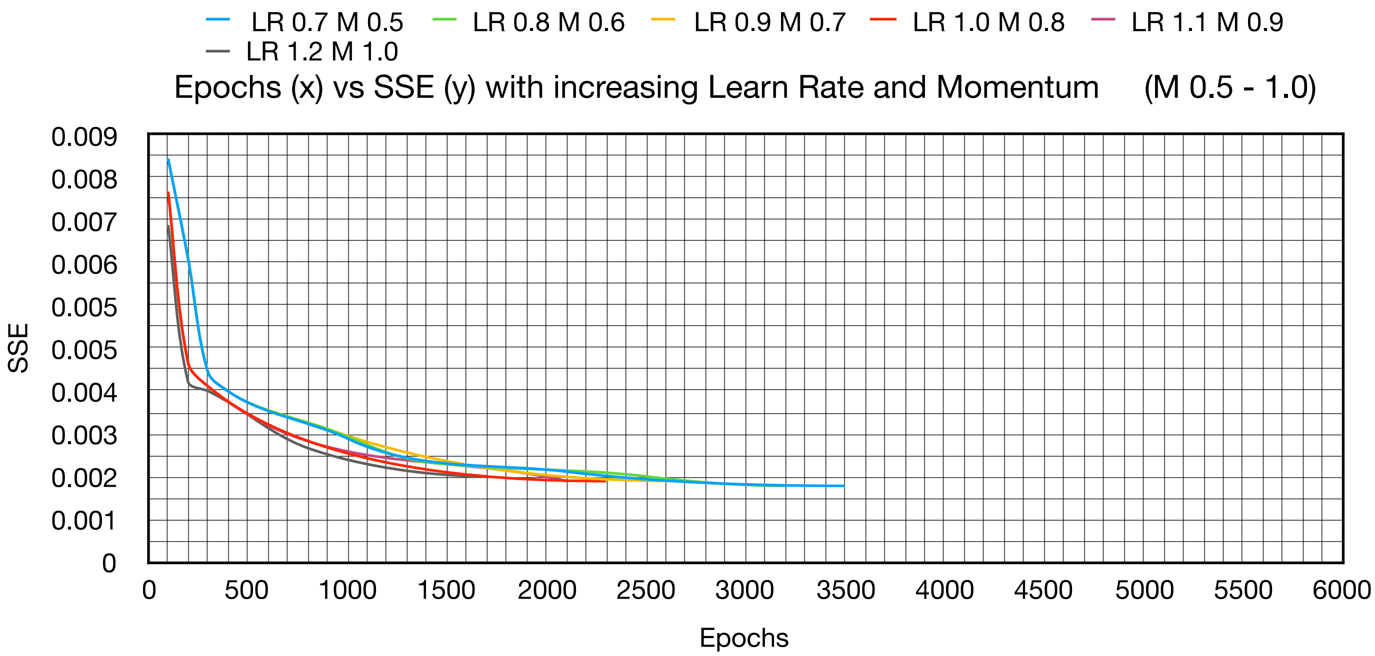
The second experiment considered an increasing momentum. The momentum increases from 0.0 to 1.0 in increments of 0.1. Without a change in learning rate, a change in momentum will not affect the results, as such the learning rate is also changed, increasing from 0.2 to 1.2 at increments of 0.1. For clarity between lines on the graphs, the results have been split into two graphs.

The first graph shows an overall trend, as the learn rate and the momentum increase, the number of epochs decreases. The final SSE is found within a range of 0.0015 to 0.0020. The number of epochs decreases from 5760 to 1960 at a learning rate of 0.5 and momentum of 0.3. The final SSE is lowest at 0.0015 after 4490 epochs at a learning rate of 0.6 and momentum of 0.4.

The second graph shows the same overall trend as the first graph, as the learn rate and the momentum increase, the number of epochs decreases. The final SSE is found within a range of 0.0016 to 0.0018. The number of epochs decreases from 3560 to 1720 at a learning rate of 1.2 and momentum of 1.0, this is the lowest number of epochs however it does not have the lowest SSE.

The overall range of final SSE values is 0.0015 to 0.0020. The gradient of the second graph is steeper than the first graph, showing a greater change in SSE values as the learning rate increases. Furthermore the greatest change in SSE is found within the first 500 epochs. Learn rate is not directly proportional to the number of epochs.





4.3 Experiment Three: Hidden Neurons

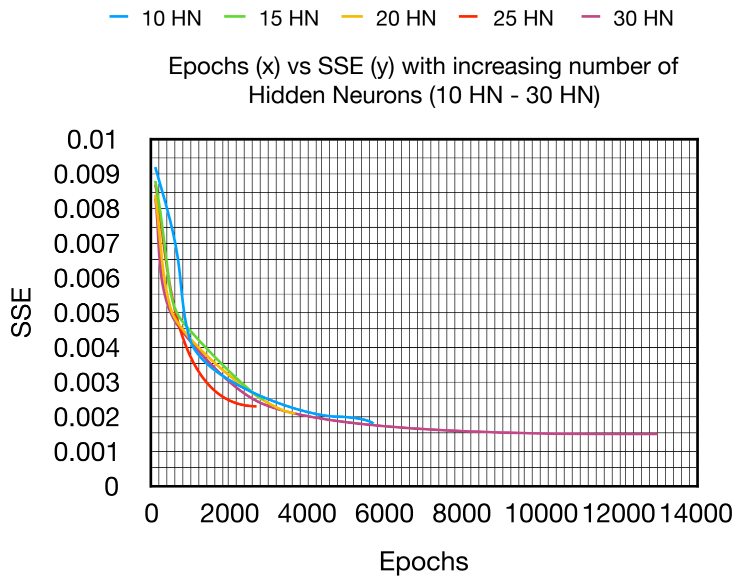
The third experiment considered an increasing number of hidden neurons. The network initially had 10 hidden neurons, increasing to 60 in increments of 5.

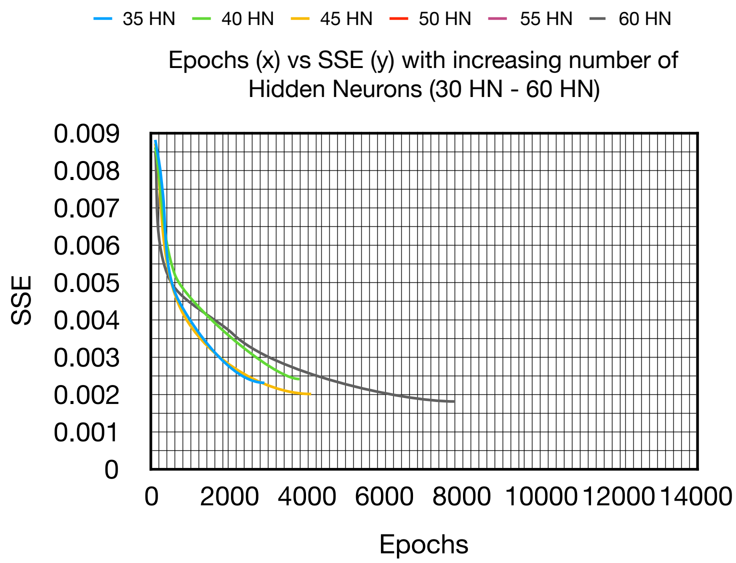
The two graphs do not show any overall trends related to an increase in the number of hidden neurons. The number of epochs does not decrease as the number of hidden neurons increases.

The second graph omits results for 50 and 55 hidden neurons. They reached final SSEs of 0.0092 a d 0.0091 after 70 epochs and 30 epochs respectively. This shows the network could not reduce the error and thus stopped learning.

The second graph shows a higher final SSE (0.0023) and a greater number of epochs (3870) at 40 hidden neurons than at 35 hidden neurons (0.0022 and 2980).

The rate of change in SSE is greatest within the first 1000 epochs. The first and second graphs have similarly shaped curves, showing little overall variation.





5. Discussion of Results

The first experiment clearly shows as the learning rate increases the number of epochs required to reach the final SSE decreases. The number of epochs decreases between learning rates of 0.2 and 0.5 but increases at 0.6. Between learning rates of 0.6 and 1.2 it once again decreases. This suggests that the best course of action to decrease the time required to learn is not just to increase the learning rate. The range in SSE is from 0.0015 to 0.0020, this proportionally much lower than the range in number of epochs required. This suggests a greater learning rate will not substantially affect the final SSE, it has a greater effect on the number of epochs required to reach this final value.

The second experiment offers the same numerical results as the first experiment. This shows that in learning the colour codes, momentum does not have any effect on the results, as such it does not need to be considered when deciding upon the best parameters for the neural network.

The greatest change in SSE is found within the first 500 epochs, the results then curve out. This suggests that if constrained for time, it may be advantageous to reduce the maximum number of epochs as after a certain point the decrease in SSE may be considered too little to justify the time spent learning.

The third experiment shows there is no strong relation between the number of hidden neurons and the number of epochs required to reach the final SSE. The results for 40 hidden neurons and 35 hidden neurons suggest for certain parameters the network will be unable to the learn the colour codes without a large amount of error, rendering it ultimately unreliable for suitable application.

6. Conclusion

A greater learning rate will not substantially affect the final SSE, it has a greater effect on the number of epochs required to reach this final value.

In learning the colour codes, momentum does not have any effect on the results.

It may be advantageous to reduce the maximum number of epochs as after a certain point the decrease in SSE may be considered too little to justify the time spent learning.

There is no strong relation between the number of hidden neurons and the number of epochs required to reach the final SSE.

References:

1. https://www.investopedia.com/terms/n/neuralnetwork.asp Date accessed: 26/02/20
2. https://machinelearningmastery.com/learning-rate-for-deep-learning-neural-networks/ Accessed 26/02/20
3. https://stats.stackexchange.com/questions/338255/what-is-effect-of-increasing-number-of-hidden-layers-in-a-feed-forward-nn Date accessed: 26/02/20

4. Brownlee, Jason (2017-07-13). "What is the Difference Between Test and Validation Datasets?

5. Ripley, Brian (1996). Pattern Recognition and Neural Networks. Cambridge University Press. p. 354

6. https://htmlcolorcodes.com/color-names/ Date accessed 26/02/20