ESI 4244 Project Report

2^k Analysis of Image Recognition Neural Network

Vernicia Gage

Marcus Northrup

Executive Summary

A 2^k factorial design was proposed to assess a neural network but due to time constraints, a ¼ fraction factorial analysis was done instead. High and low values for each hyperparameter was established and experimentation and calculations were done to determine what factors and interactions had a significant impact on the system and to and to find a combination of factors that would result in a validation accuracy within the anticipated values of 50 – 60%.

It was concluded that while a fractional 2^k factorial analysis can be performed and conclusions can be made about the linearity of the system and a "satisfactory" resulting validation accuracy can be obtained through trial and error during experimentation, more through experimentation should be done to shed more insight on this system. Due to the apparent non-linearity of the system, a 3-factor experimental design is recommended next to assess if there are quadratic effects from factors and interactions in the system.

Introduction

In biology, there exists the concept that biological systems can be represented as networks that contain interactions and relations between entities(4). These biological networks are not unlike a process flowchart, where steps are executed based on entity interactions and relations. One such network is the collection of neural circuits in the brain. The idea is that the brain " is assembled from local neural circuits that are connected into networks"(2). Information is distributed between neurons and the brain makes classifications and decisions based on such information. This idea is then transferred over to [Artificial] Neural Networks. This neural network is comprised of "simple processing nodes" (1) that contain layers that takes in initial input data, processes it through the layers, making decisions to define the data, then after being processed through the final layer, returns the output. Neural networks are "first trained having weights and thresholds set to random values"(1) which data is fed through it over and over with weights being "continually adjusted until training data with the same labels consistently yield similar outputs"(1).

The modern concept of Neural networks was put forth in 1943 by Warren McCulloch and Walter Pitts, who introduced the concept of artificial [network] neurons.(5) Since then the study and use of Neural Networks has progressed to predicting stocks, facial recognition, image processing and even predicting tool life of machines(6).

With the usefulness of a theoretical perfect neural network knowing no bounds, the problems that arise are 1) how to improve the accuracy of a neural network and 2) how conduct the process from start to finish, efficiently. This experiment will focus on the prior, how does one improve the accuracy of a neural network.

Problem Description

The usefulness of a neural network is capped by the accuracy of its prediction. No matter how fast it processes data or how little training it requires, an inaccurate network is a useless one. In each neural network there can be tens of hundreds of thousands of parameters that need to be defined and adjusted to produce the highest prediction accuracy, the calculation of which by hand would be tedious and exhaustive and even by computer would be long and drawn out if all combinations of all factors needed to trained and tested. To save time and effort, a 2^k factorial design will be used to help determine the best combination of hyperparameters to produce the highest level of accuracy.

Project Objective

The objective of this experiment is to conduct a 2^k factor analysis on the hyperparameters defined in the image processing neural network given and find the “best combination of hyperparameters that gives the highest validation accuracy in a neural network” (7).

Methodology

This experiment has 9 factors, and we are choosing to analyze it using a high and low level for each hyperparameter, this makes this experiment a 2^9 factorial design experiment. A 2^9 factorial design has 512 experiments to perform. As time is constrained and this is a preliminary experiment, a faster method for data collection was used at the cost of accuracy. It was decided to conduct a fractional factorial experiment instead, this reduces the number of experiments that are necessary to perform. A 2^{9-2} or 2^7 experiment was conducted. This is a ¼ fraction factorial analysis, so we only had to conduct 128 experiments confounded the other ¾ of the interactions with the ¼ used in the experiment. From here, we conducted the 128 experiments.

Experiment and Data collection

The first step in conducting the experiment was to run the program at the default levels given. From there each hyperparameter was changed to various values within acceptable bounds while keeping the remaining hyperparameters at their default values. The graphed output of which are in *Appendix 1*. A record of the value of the hyperparameter value that produced the lowest and the highest validation accuracy (*see Figure 1*) was kept and that was used for the next step in the experiment.

Parameter	Default	Low	High	Best Output
Kernel Size	5	1	11	11
Convolutional Layers	1	4	1	1
Num. of Filters	64	16	264	264
Activation	sigmoid	sigmoid	elu	sigmoid
Neurons	300	100	200	100
Learning Rate	0.001	1	0.10	0.10
Num. of Epochs	10	1	40	40
Batch Size	10	100	1	100
Momentum	0.80	0.50	0.99	0.50

Figure 1: Table of hyperparameter values at the default, the individual low and high hyperparameter values and the experimental combination that yielded the best output.

Step two is to perform the ¼ factorial design; We confounded I = ACDFGH = BCEFGI = ABDEHI in this design. From there we ran the neural network with the decided upon experimental combinations and obtained our data.

Data Analysis

After data was collected, a normality plot of the data was constructed (*see Figure 2*).

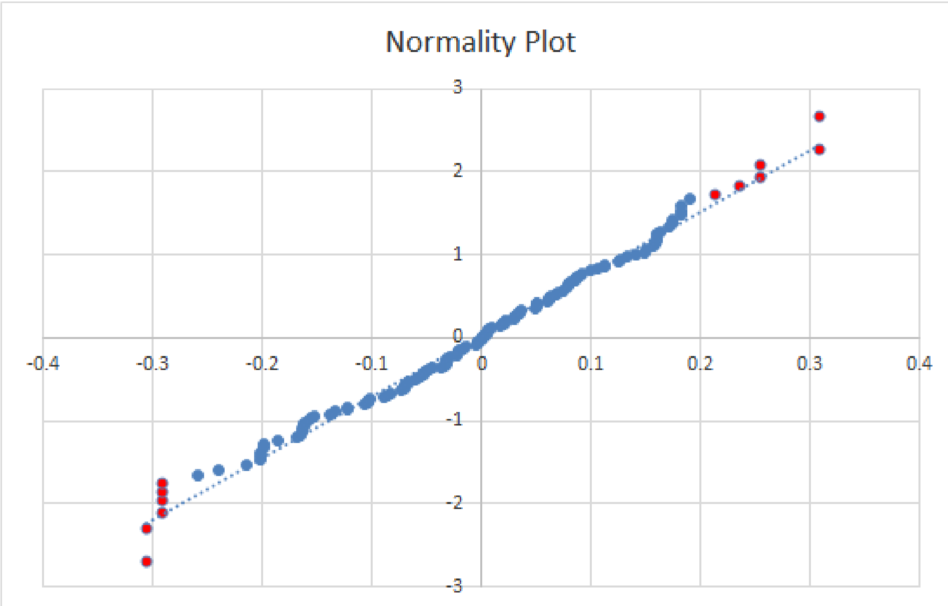


Figure 2: Normality plot of the resulting data, with outliers highlighted in red.

Through visual inspection of the normality plot, it was determined that there were 12 outliers. From there, the f_0 values were compared to the $f_{critical}$ value with an alpha value of $\alpha = 0.001$ to leave us with four significant factors and three significant interactions.

Results

Through fractional factorial analysis (2^{9-2}) it was determined that the significant factors are: B, F, H and I and that the significant interactions are AGI, BDF, and ABCDEHI. Our β_0 value was calculated to be 0.129. This leads us to the linear regression equation:

$$y = 0.129+0.126B+0.154F-0.146H-0.153I-0.129AGI-0.120BDF-0.146ABCDEHI$$

Trial and error experimentation with these parameters lead us to receive a validation accuracy of 56.67% while only using the high and low values of the hyperparameters.

Concluding Remarks

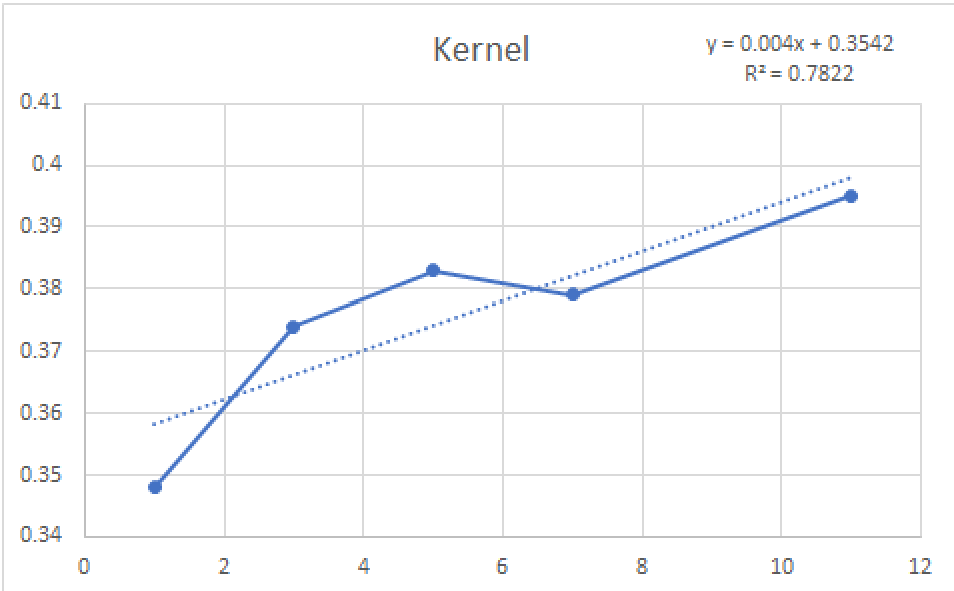
Though a validation accuracy was obtained that was within acceptable bounds of this experiment, there are improvements that can be made to the experimental design to improve on the regression equation and give more credibility to the accuracy of the values obtained in this experiment. The full 2^9 factorial design can be conducted to get more accurate data by eliminating confounding, which makes us unable to tell whether a factor or interaction is causing a significant impact on the experiment, or whether the impact is due to the variables or interactions confounded with it. Conducting the experiment multiple times will help to obtain higher accuracy for the effect of each hyperparameter and their interactions by giving more degrees of freedom of the error, resulting in accurate values for the impact of the factors and interactions. Finally, this problem is likely more complex than a 2k factorial design can describe, and in further conducting this experiment, a 3-level factorial design should be considered to generate quadratic models.

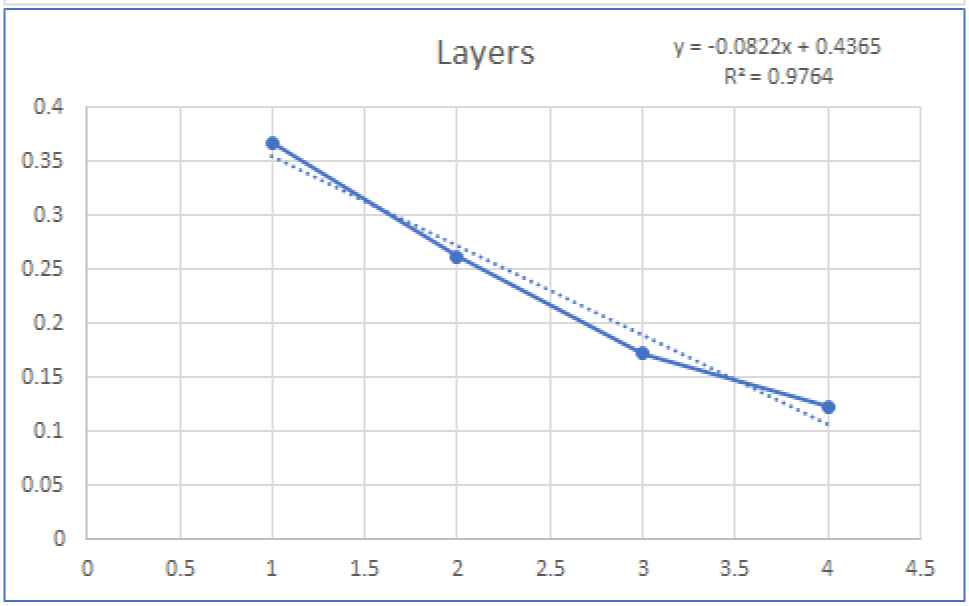
References

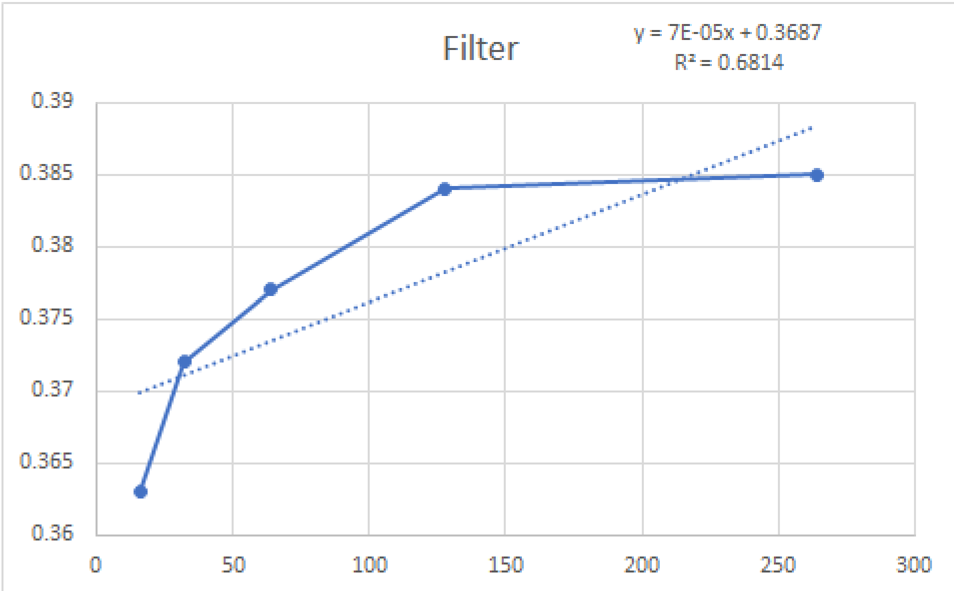
1. https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414
2. https://www.brainfunction.edu.au/research/research-themes/neural-circuits/
3. https://towardsdatascience.com/understanding-neural-networks-19020b758230
4. https://www.ebi.ac.uk/training-beta/online/courses/network-analysis-of-protein-interaction-data-an-introduction/network-analysis-in-biology/
5. Piccinini, G. The First Computational Theory of Mind and Brain: A Close Look at Mcculloch and Pitts's “Logical Calculus of Ideas Immanent in Nervous Activity”. Synthese 141, 175–215 (2004). https://doi.org/10.1023/B:SYNT.0000043018.52445.3e
6. https://www.sciencedirect.com/science/article/pii/S0898122110002555
7. Hints for Doe Project, document provided for ESI 4244 Project, Fall 2020

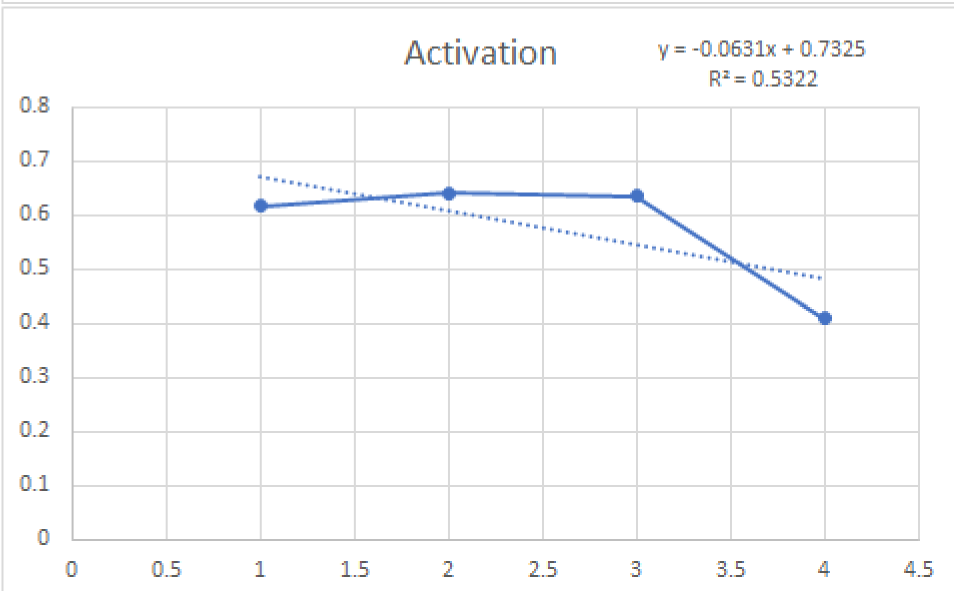
Appendix

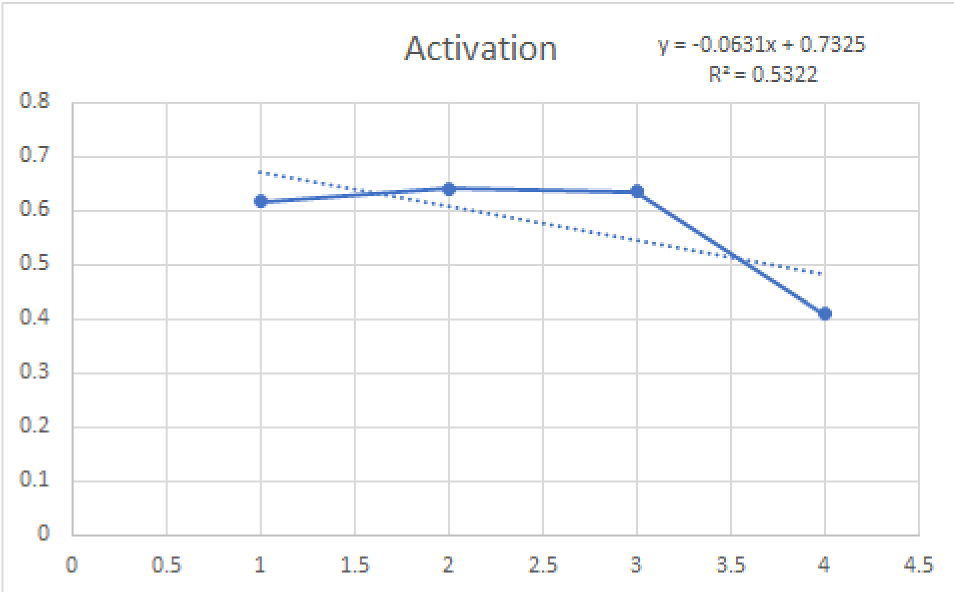
Appendix A: Graphs of Hyperparameter Effects on the System

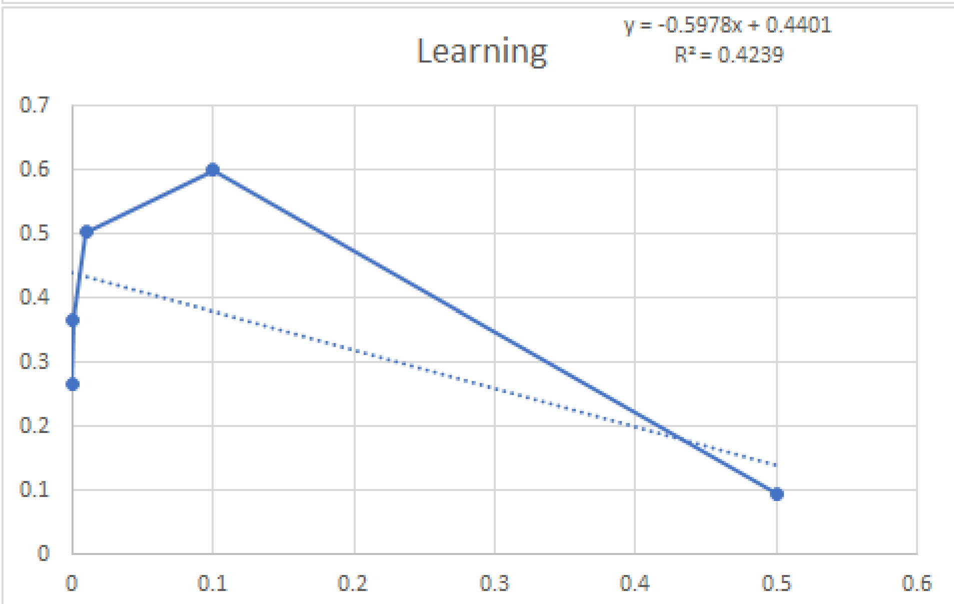


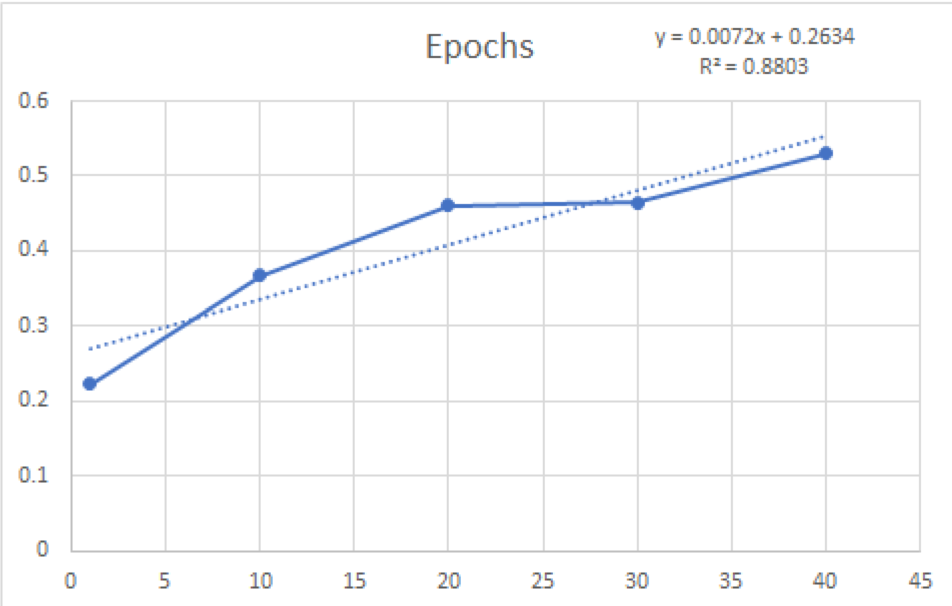


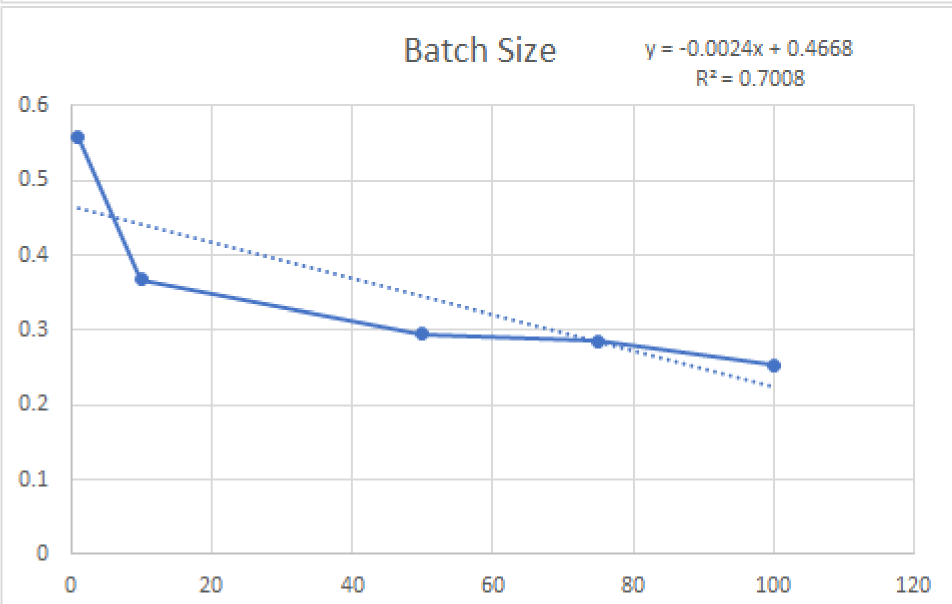

Filter
y = 7E-05x + 0.3687
R² = 0.6814
0.39
0.385
0.38
0.375
0.37
0.365
0.36
0
50
100
150
200
250
300

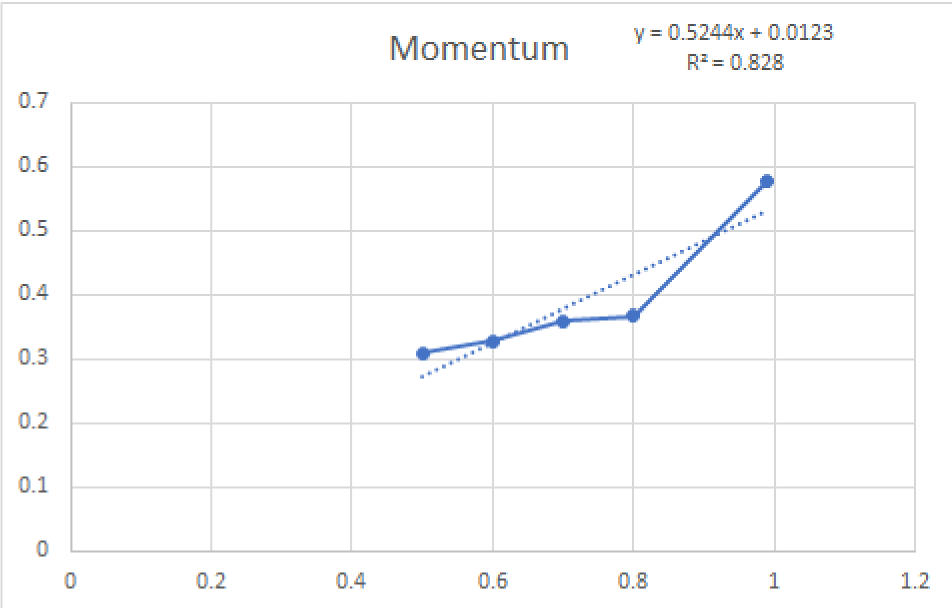

Activation
y = -0.0631x + 0.7325
R² = 0.5322
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5


Activation
y = -0.0631x + 0.7325
R² = 0.5322
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5


Learning
y = -0.5978x + 0.4401
R² = 0.4239
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6


Epochs
y = 0.0072x + 0.2634
R² = 0.8803
0.6
0.5
0.4
0.3
0.2
0.1
0
0
5
10
15
20
25
30
35
40
45


Batch Size
y = -0.0024x + 0.4668
R² = 0.7008
0.6
0.5
0.4
0.3
0.2
0.1
0
0
20
40
60
80
100
120



Appendix B: Normality Plot with Highlighted Outliers