TAFFD's Journal
ISSN 2643-4938 (online) · ISSN 2769-0490 (print)

OPEN ACCESS

Effect of Selected Hyper Parameters on Accuracy Level of Deep Belief Network Algorithm Based Web Phishing Detection System

Chigozie-Okwum Chioma Chinyere & Ajah Ifeyinwa Angela

Published in TAFFD's Journal, Vol. 2, Issue 2, pp. 12–18, 2020
Article page https://journal.taffds.org/articles/effect-of-selected-hyper-parameters-on-accuracy-level-of-deep-belief-network-algorithm-based-web-phishing-detection-system-2-12

CC BY 4.0 **Licence**

© 2020 The Author(s). Published by Transdisciplinary Agora for Future Discussions Inc. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Licence terms: https://creativecommons.org/licenses/by/4.0/

This page was added by the journal. The article begins on the next page.

Effect of Selected Hyper Parameters on Accuracy Level of Deep Belief Network Algorithm Based Web Phishing Detection System.

Chigozie-Okwum Chioma Chinyere[1]& Ajah Ifeyinwa Angela[2]
[1]Spiritan University Nneochi. [2]Ebonyi State University, Abakaliki.
Email Id: chiomaokwum@gmail.com[1]; ifyajah@gmail.com[2].

Abstract

The research sought to investigate the effects of altering values of three selected hyperparameters (Epoch, Learning rate and Drop out), on the accuracy level of a deep belief network algorithm based web phishing detection system. The accuracy level of the system changed as these optimiser values were altered. The experimental results showed that there was a positive linear relationship between Epoch and learning rate values with regards to the accuracy of the system. The study identified a negative linear relationship between dropout values and accuracy of the web phishing detection system. The study hence recommended that keeping Epoch values high around or a little above 25 cycles, maintaining a lowdrop out value (0.1), while increasing the value of learning rate (>=0.7) would ensure arrival at a high accuracy level of the deep belief network algorithm based web phishing detection system (0.987).

Keywords: Hyper Parameters, Epoch, Dropout, Learning Rate, Machine Learning.

Introduction

The growing need for more complex and computationally expensive machine learning models has resulted in tremendous growth in research on hyperparameters and hyperparameter optimisation. Machine learning research in recent years is focused on ensuring automation and reduction in human interference. The help of hyperparameters achieves the objective; these optimisers ensure peak performance of the model when tuned correctly. Every machine learning model possesses hyperparameters and automation to ensure optimal tuning of these parameters to achieve the best results. According to Fuerer & Hutter, (2019), regularisation, optimisation and network architecture of recent deep neural network algorithms are dependent on an array of hyperparameters. Varying the value of hyperparameters have a reflection on the performance of a model.

The Problem & Objective

According to Jan et al. (2018), it is usually unknown, which hyperparameter to tune to achieve optimal system performance when testing out a new algorithm on a data set. Furthermore, it is salient to determine what suitable ranges are for the optimisers and

which values in the range would produce a high performance of the system. Well-known algorithms like the support vector machines, it is already established as to which optimiser has a significant effect on the performance of the model, gamma and complexity optimisers are the parameters that affect the performance of the model the most. The motivation for this work hence is to identify the critical influencing optimisers for a deep believe network algorithm based web phishing detection model.

In this work, the deep belief network algorithm was used in developing a web phishing detection system, and an attempt is made towards identifying the effects of selected hyperparameters on the performance of the system and at what value the selected hyperparameter is optimised. The hyperparameters considered include Epoch, Learning rates, and Drop out values. The paper introduces the concept of hyperparameter optimisation, reviews related works supporting the optimisation of selected hyperparameters, identifies the procedure of the experiment, discusses results of the experiments and thus arrives at the identified conclusion.

Research Questions

The following research questions were answered in the course of the study.

1. What is the effect of epoch value on the accuracy of a Deep believe network algorithm based web phishing detection system?
2. What is the effect of Learning rate on the accuracy of a Deep believe network algorithm based web phishing detection system?
3. What is the effect of drop- out value on the accuracy of a Deep believe network algorithm based web phishing detection system.

Literature Review

Hyperparameters are settings that influence how well a model trained when they are tuned; they are generally set before the training process begin. Rodrigues (2018) supports that although existing outside a model, there is a correlation between parameter and accuracy of the model. Hyperparameters affect the model positively or negatively. Some examples of hyperparameters in machine learning are Learning Rate, Number of Epochs, Momentum, Regularisation constant, Number of branches in a decision tree, number of clusters in a clustering algorithm (like k-means), dropout, Hidden layer structure among others.

Epoch value is the hyperparameter that determines how many cycles the algorithm will transverse through the training data set. Having the model cycle once usually causes overfitting and running more epoch help balance the fitting. Increasing the epoch value

changes the model from an overfit to an optimal fit but should be watched to avoid arriving at an underfit. (Brownlee, 2018).

Learning rate is an optimiser that helps to determine to what extent the machine learning algorithm can be changed with regards to an estimated error at every time there is an update in the weight of the model. It has a tremendous impact on the model and should be considered critically. Learning values range between 0.0 and 1.0. Choosing the learning rate is challenging as a value too small may result in a lengthy training process that could get stuck, whereas a value too large may result in learning a sub-optimal set of weights too fast or an unstable training process. (Brownlee, 2019).

Dropout is the optimiser in a machine learning model that helps to avoid overfitting in the model. Dropout achieves its aim of reducing overfitting by setting the outgoing edges of hidden units (neurons that make up hidden layers) to point of 0(zero) at each update of the training phase randomly (Maklin, 2019). Dropout permits neurons to be ignored during the training of a group of neurons randomly selected. It means that these ignored units referred to as neurons are not considered during a specific forward or backward pass. Dropout refers to ignoring units (i.e. neurons) during the training phase of a specific set of neurons which is chosen at random. Ignoring Units mean these units are not considered during a particular forward or backward pass.

According to Rodrigues (2018), hyperparameters are very model specific. Optimising hyperparameters is very beneficial in a machine learning model as it ensures minimal human effort, they ensure that a model is trained to be problem-specific when optimised help reduce human efforts in applying machine learning to tasks. (Melis, et al., 2018; Snoek, et al., 2018). Hyperparameters also improve the ability to reproduce several results, and this helps in aiding ease and fairness of research results. Hyperparameters ensure a balanced comparison; this is because the fair comparison is possible only when the same level of tuning for the problem at hand is achieved.

Methodology

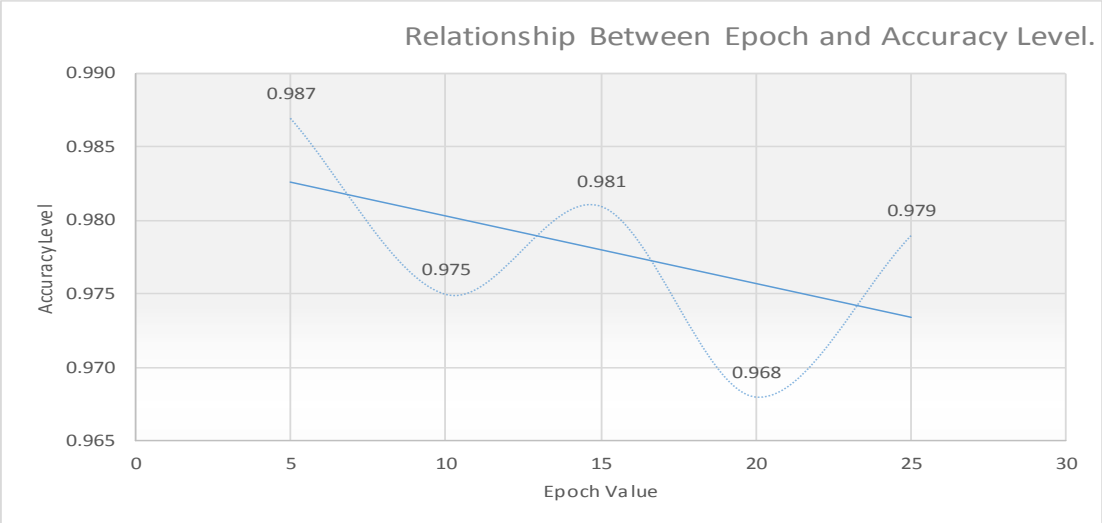
The experimental research method was used in this study to provide solutions to research problems. The study adopted the experimental research method because it allows the researchers to alter settings and tunings of the model and observe the changes thereof. Experimentation enabled the researchers to identify changes in the behaviour of the model when different hyperparameters were altered. In this study, the effect of the selected hyperparameters on the accuracy level of a Deep belief network Algorithm-based web phishing detection system was also identified. The experimental procedure commenced

with the development of the system. The system was developed using Python programing language and deploying a deep belief network algorithm. The web phishing detection system was developed in the following four steps: pre-processing, feature selection and extraction, model development, and prediction. In order to ascertain the effects of the learning rate, epoch values and drop out values, an experimental environment was set up on Jupyter notebook. This is an environment used in running data science and machine learning experiments. After setting up the experimental environment, the values of web phishing began to tune the values of Epoch, learning rate and drop out. We studied how changes to these selected hyperparameters affected the accuracy level of the system. The changes identified were logged and plotted out on the graph. It enabled the researchers to identify the relationship between each optimiser and the accuracy level of the web phishing detection system.

Results

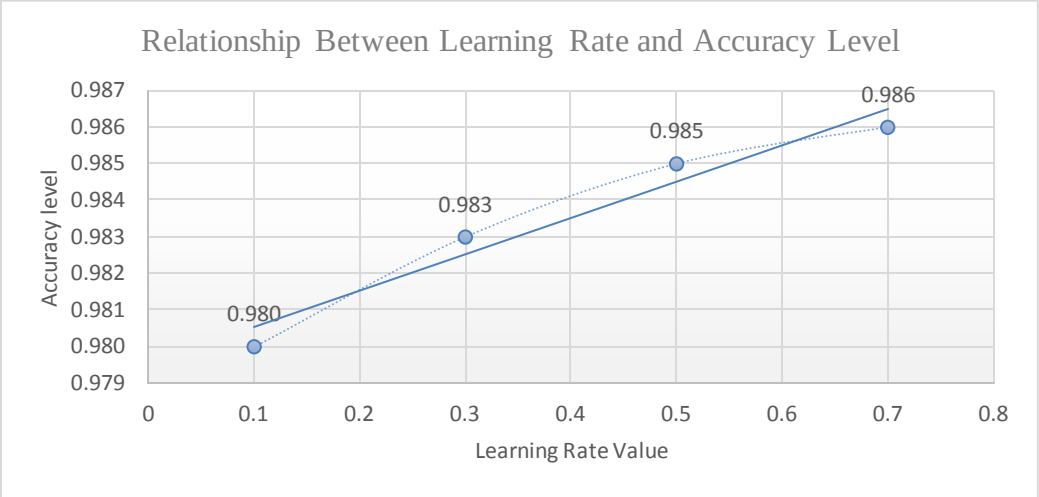
Experimental data were analysed using graphs, and the results are presented below:

Figure 1 Epoch Values and the accuracy of the DBN-based web phishing detection system.



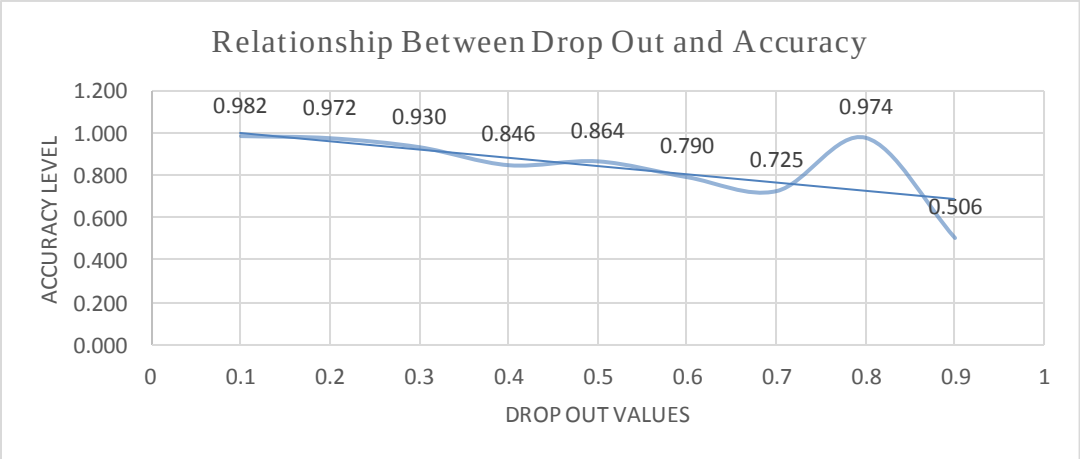
The experimental results, as shown in the graph in figure1above, outlines the estimated and actual model. The actual model is the curve with the crests and curves while the estimated model is the line of best fit (the straight-line graph). The line of best fit from the graph shows the linear relationship between the two variables (epoch value and accuracy level). The result shows accuracy level picking up at 25 epoch cycles hence depicting a positive linear relationship between epoch values and accuracy levels of a DBN-based web phishing detection system. This positive relationship implies that the accuracy level of the web detection system increased with increase in epoch values. It also means that the higher the number of iterations, the more likely to have an increase in accuracy levels.

Figure 2: Learning Rates and the accuracy of the DBN-based web phishing detection system.



The result, as seen on the graph in figure 2 above, shows a positive linear relationship between learning rates and the accuracy of the deep believe network algorithm based web phishing detection system. The results show that an increase in learning rates produces an increase in the accuracy of the model. When learning rates increased to 0.7, the accuracy level rose to 98.6%. Consequently, lower learning rates of 0.1 produces a lower accuracy level of 98%.

Figure 3: Relationship between Drop out and the accuracy of the DBN-based web phishing detection system.



The results as seen in the graph on figure 3 above presents a negatively sloped graph. It hence shows a negative linear relationship between drop-out values and accuracy levels of a deep belief network algorithm based web phishing detection system. With an increase in drop out values, the accuracy of the model reduced significantly. The dropout values of 0.1 gave an accuracy level of 98.2%, while dropout value of 0.7 pushed down the accuracy level of the system to 50.6%.

Discussion of Findings

The findings of the experimental research as seen in the charts as results presented shows that increasing the number of cycles/epoch value the model traverses through the data set. It tends to improve the fitness of the model hence increasing its accuracy. This finding is in line with (Brownlee, 2018), which suggests that increasing epoch value improves a model from an overfit to an optimal model. The study recommends leaving epoch values at around 25 cycles to avoid turning the model into an underfit.

Furthermore, the study identified a positive linear relationship between the learning rate and accuracy level. The model achieved a high accuracy level of 98.6%; when learning rate value was increased to 0.7. Learning rate values of 0.1 achieved lower accuracy rates for the model. The value of the learning rate should be chosen carefully to ensure optimisation. Finally, the study revealed that there was a negative linear relationship between drop out values and accuracy of the model. Increased drop out values reduced the accuracy level. The study hence recommends selecting epoch values of 25 cycles, learning rate values of 0.7 and a dropout value of 0.1 as the optimisation stage, at which the model performed at its best accuracy level.

Conclusion & Suggestion for Future Work

The study identified a positive linear relationship between epoch values and accuracy level of the deep belief network algorithm based web phishing detection system. It further identified a negative relationship between dropout values concerning the accuracy of the web phishing detection system. Increase in drop out values caused a decrease in accuracy levels of the system. However, there was a positive linear relationship between learning rate values and the accuracy of the DBN-based web phishing detection system. Increase in learning rate values caused an increase in accuracy values of the system.

The study showed that out of the three hyperparameters (Epoch, Learning rates and Dropout) examined in the experiment, changes in drop out had the most significant effect on the accuracy level of the model. Changes in dropout values have a drastic reducing effect on the accuracy of a web phishing detection system, reducing it significantly from 98.2% to 50.6%.

The study therefore recommended that keeping Epoch values high around 25 cycles, and maintaining a low dropout value (0.1), while increasing the value of learning rate (>=0.7) would ensure arrival at a high accuracy level of the deep belief network algorithm based web phishing detection system (0.987). Identifying appropriate technique to apply in optimising these hyperparameters is a suggestion for future work.

References.

Bergstra, J., Yamins, D. & Cox, D. (2014). Making a science of model search: Hyperparameter optimisation in hundreds of dimensions for vision architectures. In Dasgupta & McAllester (Ed.), Proceedings of the 30th International Conference on Machine Learning (ICML'13). Omnipress, 115–123.

Brownlee, J. (2018). Difference between a batch and an epoch in a neural network. Retrieved from https://machinelearningmastery.com/difference-between-a-batch-and-an-epoch/.

Brownlee, J. (2019). Understand the impact of learning rate on neural network performance. Retrieved from https://machinelearningmastery.com/understand-the-dynamics-of-learning-rate-on-deep-learning-neural-networks/.

Cory, M. (2019). Dropout neural network layer in keras explained. Retrieved from https://towardsdatascience.com/machine-learning-part-20-dropout-keras-layers-explained-8c9f6dc4c9ab.

Feurer. M. & Hutter. F. (2019). Hyperparameter optimization. *Machine Learning* , 3-33. doi: 10.1007/978-3-030-05318-5_1.

Jan N. van Rijn & Hutter, F. (2018). Hyperparameter importance across datasets. In KDD '18: The 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, United Kingdom. ACM, New York, NY, USA, https://doi.org/10.1145/ 3219819.3220058

Melis, G., Dyer, C. & Blunsom, P. (2018). On the state of the art of evaluation in neural language models. Proceedings of the International Conference on Learning Representations (ICLR'18). Retrieved from http://www.iclr.cc/

Rodrigues, J. (2018). Understanding hyperparameters optimisation in deep learning models: concepts and tools. Retrieved from https://towardsdatascience.com/understanding-hyperparameters-optimization-in-deep-learning-models-concepts-and-tools-357002a3338a.

Sculley, D., Snoek, J., Wiltschko, A. & Rahimi, A. (2018). Winner's curse? On pace, progress, and empirical rigour. Proceedings of International Conference on Learning Representations Workshop track. Retrieved from http://www.iclr.cc.

Snoek, J., Larochelle, H. & Adams, R. (2016). Practical Bayesian optimisation of machine learning algorithms". In Bartlett et al. (Eds.), Proceedings of Machine Learning Research, 48, 2960–2968.