Journal of Artificial Intelligence Research & Advances
ISSN: 2395-6720 (Online)
Volume 6, Issue 2
www.stmjournals.com

Performance Evaluation of Autism Diagnosis Tool at Diverse Learning Rates

Vikas Khullar[1,]*, Manju Bala[2], Harjit Pal Singh[3]
[1]Research Scholar, Department of Computer Science and Engineering, IK Gujral Punjab Technical University, Kapurthala, Punjab, India
[2]Research Guide, Department of Computer Science and Engineering, IK Gujral Punjab Technical University, Kapurthala, Punjab, India
[3]Research Guide, Department of Applied Science, IK Gujral Punjab Technical University, Kapurthala, Punjab, India

Abstract
Artificial Neural Networks (ANN) has been considered as a major artifact in the scenario of machine learning, which needs optimization for more accurate and efficient results. The main aim of this study is to identify the optimization techniques and their learning rates for ANN with better results. In this study, we have been utilized "Keras" library for framing ANN with different optimization techniques. The evaluation of the optimization techniques has been conducted on the basis of accuracy and loss as parameters. On the basis of our comparative study, we have proposed the best ANN optimization technique out of our evaluated techniques for the diagnosis of available Autism Spectrum Disorder (ASD) dataset.

Keywords: *Artificial Neural Network, optimization, autism spectrum disorder*

****Author for Correspondence*** E-mail: vikas.khullar@gmail.com

INTRODUCTION

Autism Spectrum Disorders (ASD) is diagnosed based on symptoms, such as qualitative impairment in social interaction, communication, restricted, repetitive and stereotyped pattern of behavior. Accurate identification of ASD is in the high requirement of clinical practitioners and parents [1–4]. Nowadays, Artificial Neural Networks (ANN) is playing the vital role in medical diagnosis systems. In comparison to human evaluation, ANN has the capability to perform certain tasks with more permanent, consistent, less expensive, and ease of duplication-dissemination with better accuracy [5].

Initially, authors had classified the various autism stages using Artificial Neural Networks (ANN), which is capable of learning and testing using a back-propagation algorithm [6, 7]. MLP neural network and logistic regression models were compared after running through a sample size of 638, showing an accuracy of not less than 92% [8]. Khullar *et al*. had collected a dataset of subjects with ASD on the basis of DSM-V criteria [5]. The authors had evaluated the collected dataset using BPNN, where they achieved nearly 100% accuracy at near 50 epochs. Loss function-based ANN had also been designed by another author who performed optimization of ANN with gradient-based and non-gradient-based methods to improve the accuracy of ANN model. Further, authors had modified the weights of ANN using optimization techniques, which also got improvement in accuracy. Chance of improvement by applying different optimizers in ANN is still available [9, 10].

The objective of the study is to identify the best performing optimizer and its learning rate for the artificial neural network for the same earlier diagnosis dataset of Autism Spectrum Disorder. This paper is divided into sections where initial section contains related studies which had been conducted to narrow down the objective. In trailing to the earlier sections, the utilized ASD dataset and ANN architecture have been discussed. Later, the graphical and tabular results have been discussed along with their detailed discussions. In the last section, the conclusion of paper has proposed the best performing optimizer and its learning rate for the artificial neural network for the diagnosis dataset of Autism Spectrum Disorder.

RELATED STUDIES

Artificial Neural Networks

For the broad range of medical diagnosis problems, ANN is being applied as a powerful tool. From the perspective of medical diagnosis, the ability of ANN to adopt continuous data allows them to solve problems that cannot be handled accurately by some conventional techniques. The basic model of ANN is composed of three main layers, i.e. input, output, and hidden layer [11]. The basic and most utilizing architecture of ANN is shown in Figure 1 [6].

Convolutional Neural Networks are recent enhancements of ANN and most commonly used types of Multi-Layer Perceptron (MLP). MLP has an activation function that maps the weighted inputs to the output which create a special structure in CNN. The CNN architecture consists of three basic layers, i.e. convolutional layer, pooling layer, and fully-connected layer, which is further connected by a rectified linear activation function for generating output [9]. The convolution operation is computed by the below equation, where y are signals, f are filters, N is the signals and x are output vectors respectively.

$$x_n = \sum_{k=0}^{N-1} y_k f_{n-k} \tag{1}$$

Optimizers

In the nutshell, an optimizer is an algorithm which minimizes the objective function at each step of ANN training by calculating the change in weights to lower the error. Keras is Python based high-level library for building neural network models where modularity and extensibility have been treated as key features [12, 13].

There is a list of optimizers available in Keras for enhancing the operational results of ANN. Here, Stochastic Gradient Descent (SGD), also known as an incremental or iterative gradient descent method has been used for optimizing a differentiable objective function. It is a stochastic approximation of the gradient descent optimization method. In the methodology, SGD stochastic gradient descent selects an observation uniformly at random.

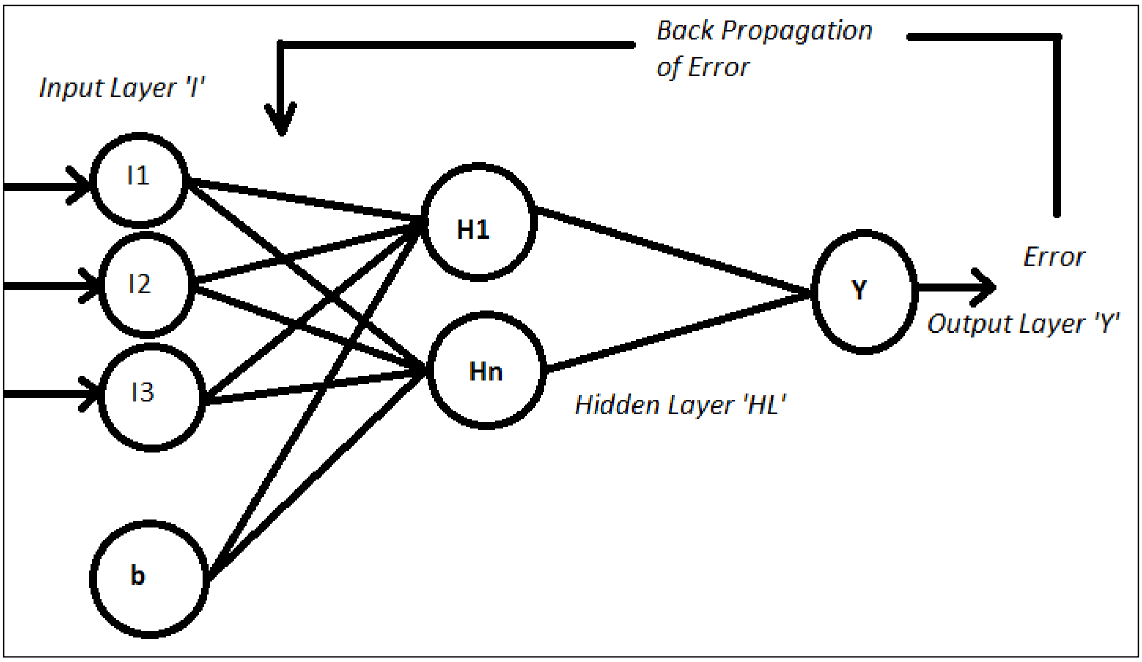

Fig. 1: *Artificial Neural Network Architecture [6].*

AdaGrad (Adaptive Gradient Algorithm) is a variation of SGD along with separately adjusted learning rate. In this method, the learning rate has been adjusted separately according to the size of the parameter gradients. This strategy improves the convergence performance over the standard stochastic gradient descent. In RMSProp (Root Mean Square Propagation), the learning rate is adapted for each of the parameters. In this method, the divided learning rate for a weight by a running average of the magnitudes of recent gradients has been used. Adadelta is related to AdaGrad optimizer. In AdaGrad working is done by accumulating all past gradients but AdaGrad uses only last N gradients. The parameters of this optimizer are similar to the ones of RMSProp. Adam (Adaptive Moment) estimation is an optimizer similar to AdaDelta and RMSProp. In addition, Adam has an exponentially decaying average of past gradients. AdaMax is nearly similar to Adam with the maximum of 2x past gradient which results from smaller and smoother movement towards the minima. Nadam (Nesterov-accelerated Adaptive Moment Estimation) is the combination of Adam and Nesterov-accelerated SGD, which results from stable movement towards minima [12–14].

Learning Rate

Learning Rate (LR) is one of the important hyper-parameters to optimize the training of ANN model. It controls the weight adjustments of our network with respect the loss gradient which travel along the downward slope. Lower learning rate ensures the addition of all local minimas [12–14]. Generally, the learning rate of a system is defined in between 0 and 1. In this study, we have used the distribution of learning rates as 0.001, 0.005, 0.01, 0.05, 0.1 and 0.5 for training and testing of optimized ANN.

$$NW = EW - LR \times G \tag{2}$$

Here, NW: New Weight, EW: Existing Weight, LR: Learning Rate, G: Gradient.

METHODS

In this section of our paper, the overall methodology has been discussed. Here, firstly we have arranged the earlier collected dataset related to the subjects with autism for training and testing. Then we have designed and applied an efficient ANN structure. Later we have been applied the different optimizers with different learning rates and analyzed the same on the basis of accuracy and loss.

Data Set for Training and Testing

DSM-V based binary dataset of 128 ASD subjects has been utilized for training and testing of optimized ANN [5].

Artificial Neural Network Structure

In this study, we had used ANN model as shown in Table 1.

This ANN architecture was trained and tested after applying different optimization techniques along with different learning rates.

Evaluation Metrics

This study has been analyzed based on parameters such as viz., accuracy, loss, mean squared error, execution time detailed as follows [8, 15]:

Accuracy

Accuracy of a medical diagnosis system may be defined as the ratio of sum of True-Positive (TP) and False-Negative (FN) with the total number of attempts (TP, FN, False Positive (FP), and True Negative (TN)).

$$\text{Accuracy} = (TP+TN)/(TP+FP+TN+FN) \tag{3}$$

Loss

Logarithmic loss (or Log Loss) is a performance metric for evaluating the predictions of probabilities of membership to a given class.

$$\text{Loss} = -(TL \times P + (1-TL) \times \log(1-P)) \tag{4}$$

TL: True Labels, and
P: Predictions.

EXPERIMENTAL RESULTS AND DISCUSSIONS

This section provides details of obtained results regarding training and testing of different optimized ANNs along with the variety of learning rates (i.e. 0.001, 0.005, 0.01, 0.05, 0.1, 0.5) as shown in the Figures 2–6.

Training Accuracy

These results reflect that in most of the discussed algorithms, with the increase in learning rate of algorithm, the quality of outcome got reduced. As shown in Figures 3–6, in comparison to other optimizers, Whereas, Adam had analyzed average initially on the basis of evaluation parameters, but later, accuracy had been reflected with the gradual improvement. Adadelta had performed with lower range of accuracy and higher range of losses. It is also noticed that, Adadelta is not providing stable results with different learning rates. It has been found that, the Adam, Adammax, and Adagrad had performed best with least losses. In the case of training and testing accuracy, all algorithms except Adadelta have provided 100% of accuracy, nearly after 10 epochs of training.

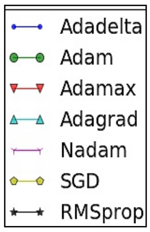

Fig. 2: *Optimizer with Representing Line.*

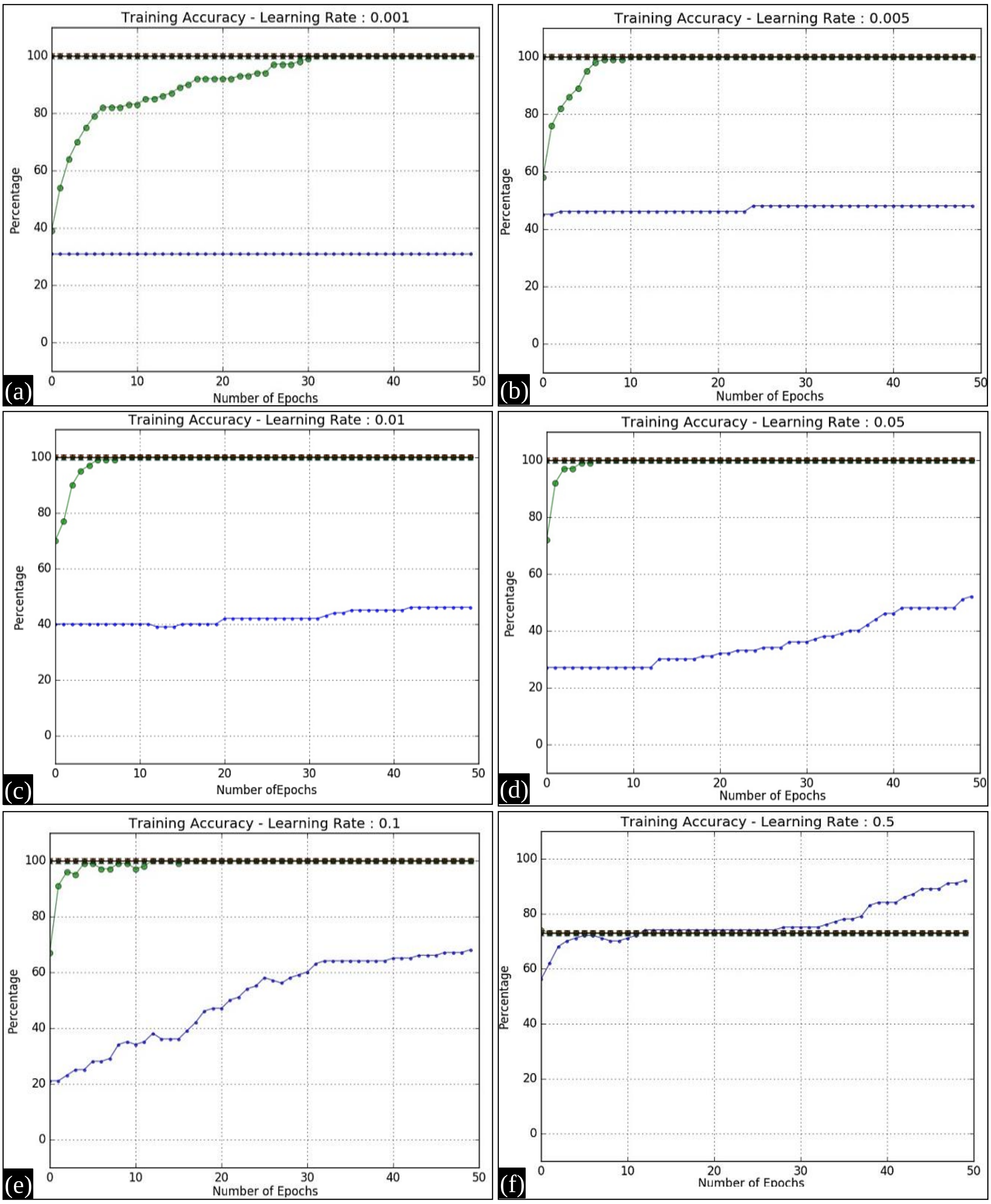

Fig. 3: *(a), (b), (c), (d), (e) and (f): Training Accuracy of Optimized ANNs with Learning Rates (i.e. 0.001, 0.005, 0.01, 0.05, 0.1, 0.5)*

Testing Accuracy

This section provides details of obtained results testing accuracy of different optimized ANN’s along with the variety of learning rates.

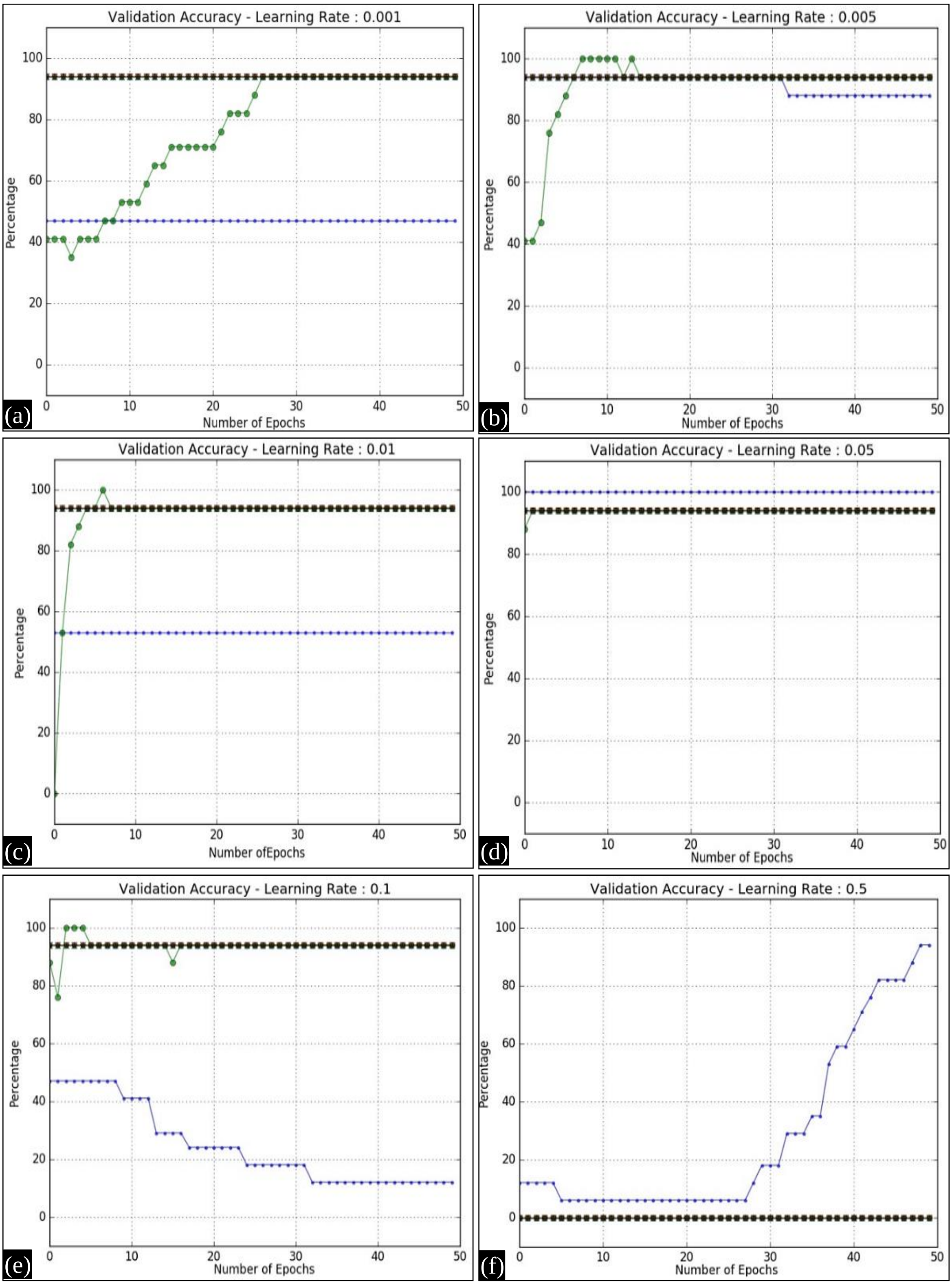

Fig. 4: *(a), (b), (c), (d), (e) and (f): Testing Accuracy of Optimized ANNs with Learning Rates (i.e. 0.001, 0.005, 0.01, 0.05, 0.1, 0.5).*

Training Loss

This section provides details of obtained results training loss of different optimized ANN's along with the variety of learning rates.

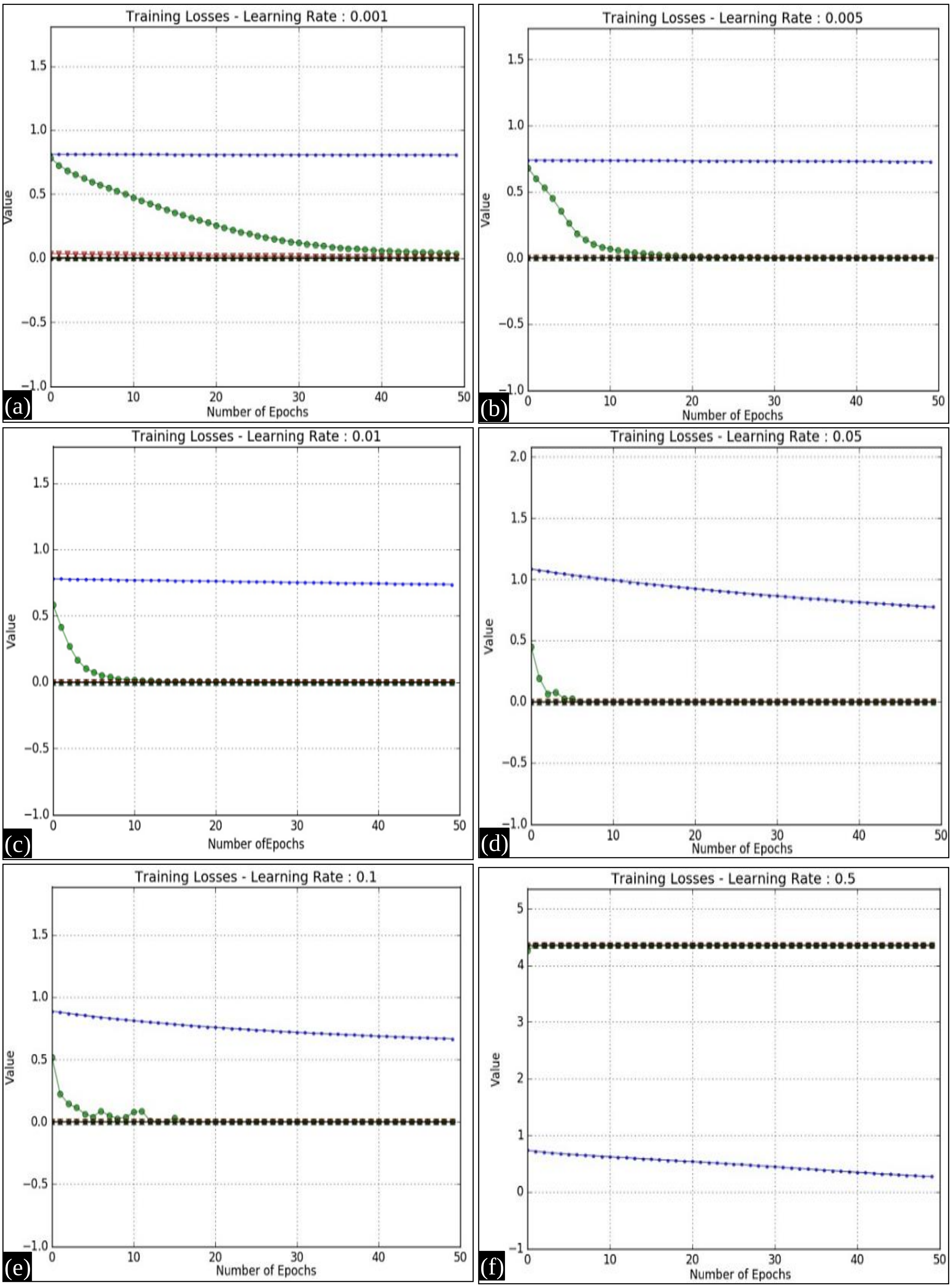

Fig. 5: *(a), (b), (c), (d), (e) and (f): Training Loss of Optimized ANNs with Learning Rates (i.e. 0.001, 0.005, 0.01, 0.05, 0.1, 0.5).*

Testing Loss

This section provides details of obtained results testing loss of different optimized ANN's along with the variety of learning rates.

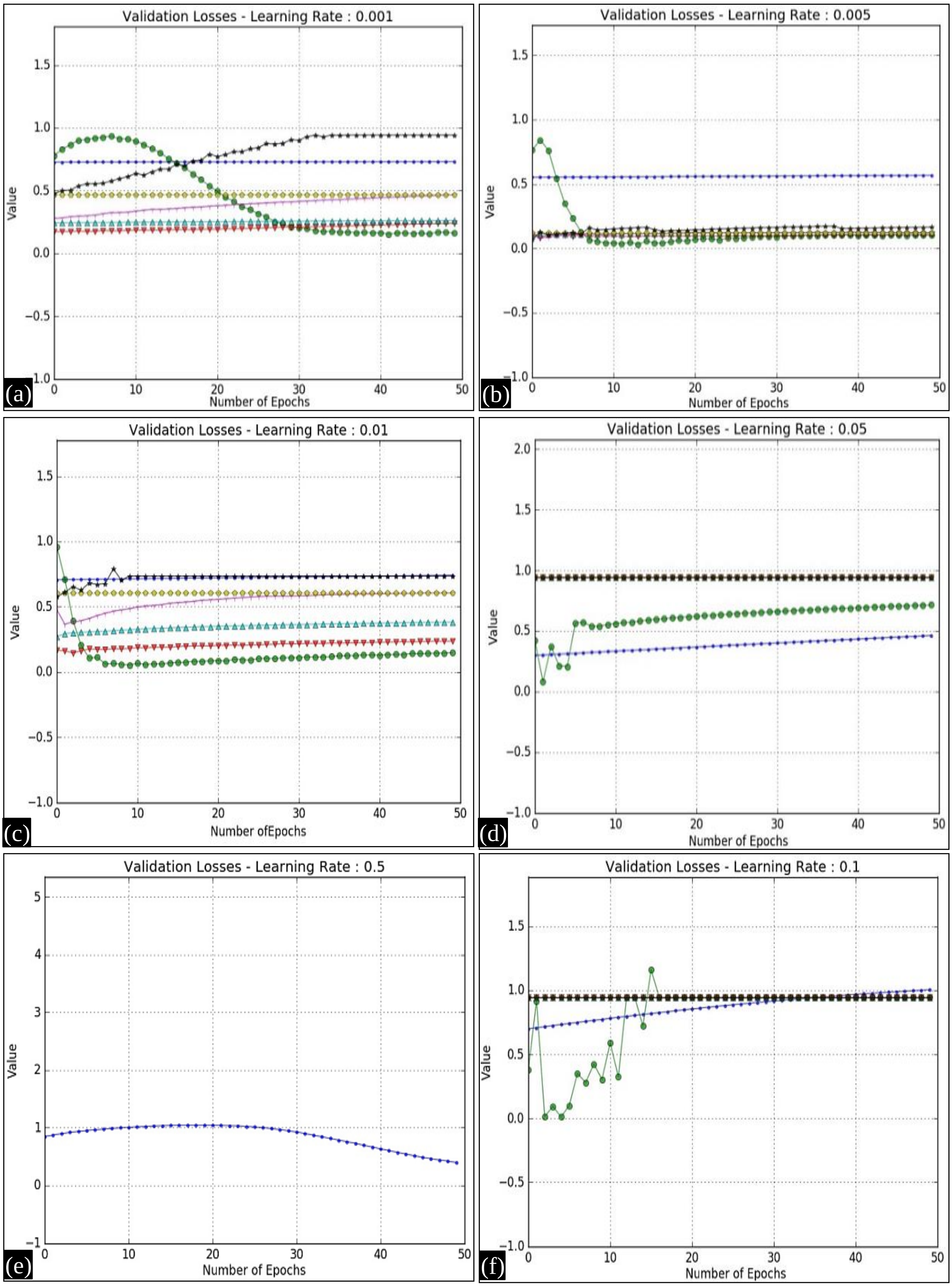

***Fig. 6:** (a), (b), (c), (d), (e) and (f): Training Accuracy of Optimized ANNs with Learning Rates (i.e. 0.001, 0.005, 0.01, 0.05, 0.1, 0.5).*

Table 1: *Structure of Utilized ANN.*

Layer Type	Output Shape	Param No.
Dense 1	(None, 7)	56
Dense 2	(None, 7)	56
Dense 3	(None, 2)	16
Activation	(None, 2)	0

CONCLUSION

From the mentioned results and discussions, we had concluded that, lower learning rate is better option for all of the ANN optimization algorithms. Further, in results of our conducted experiments, we propose the use of Adam, Adamax, and Adagrad methods for the optimization of ANN for the utilized dataset for the diagnosis of ASD.

ACKNOWLEDGEMENT

We are thankful for the I.K.G. Punjab Technical University, Kapurthala (India) and CT Institute of Engineering, Management, and Technology, Jalandhar (India) for making this research possible.

REFERENCES

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th Edn. Arlington, VA: American Psychiatric Association; 2013.
2. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders.* 4th Edn. Washington, DC: American Psychiatric Association; 1994.
3. Zwaigenbaum L, Bryson S, Rogers T, *et al.* Behavioral Manifestations of Autism in the First Year of Life. *Int J Dev Neurosci*. 2005; 23(2–3): 143–152p.
4. Mazurek MO, Kanne SM, Wodka EL. Physical Aggression in Children and Adolescents with Autism Spectrum Disorders. *Res Autism Spectr Disord*. 2013; 7(3): 455–465p.
5. Khullar V, Singh HP, Bala M. Intelligent Handheld Expert System (HES) for Diagnosis of Autism Spectrum Disorder and its Severity Level. *ASEAN Journal of Psychiatry*. 2018; 19(1): 12–26p.
6. Sudhalter V, Landon-Jimenez D, Keogh M. A Neural Network Approach to the Classification of Autism. *J Autism Dev Disord*. 1993; 23(3): 443–66p.
7. Cohen IL. An Artificial Neural Network Analogue of Learning in Autism. *Biol Psychiatry*. 1994; 36(1): 5–20p.
8. Florio T, Einfeld S, Tonge B, *et al*. Providing an Independent Second Opinion for the Diagnosis of Autism using Artificial Intelligence over the Internet. *Counseling, Psychotherapy, and Health*. 2009; 5(1): 232–248p.
9. Schmidhuber J. Learning to Control Fast-Weight Memories: An Alternative to Recurrent Nets. *Journal of Neural Comput*. 1992; 4(1): 131–139p.
10. Huang M, Wang J, Yang R. BP Neural Network Based Fault Diagnosis in Vehicle Braking Control System. *2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*. 2018.
11. Gonzalez-Fernandez I, Iglesias-Otero MA, Esteki M, *et al.* A Critical Review on the Use of Artificial Neural Networks in Olive Oil Production, Characterization and Authentication. *J Crit Rev Food Sci Nutr*. 2018; 30: 1–14p.
12. Optimizers. *Keras Documentation*. Available at Web Link: “https://keras.io/optimizers/”.
13. Ketkar N. Introduction to Keras. In: *Deep Learning with Python*. Berkeley, CA: A Press; 2017.
14. Nandy A, Biswas M. Reinforcement Learning with Keras, TensorFlow, and ChainerRL. In: *Reinforcement Learning.* Berkeley, CA: A Press; 2018.
15. Shokoohi-Yekta M, Mahmoudi M, Ghobari-Bonab B, *et al.* Developing Autism Screening Expert System (ASES). *AWER Procedia Information Technology & Computer Science*. 2013; 4: 1063–1068p.

Cite this Article

Vikas Khullar, Manju Bala, Harjit Pal Singh. Performance Evaluation of Autism Diagnosis Tool at Diverse Learning Rates. *Journal of Artificial Intelligence Research & Advances*. 2019; 6(2): 92–99p.