Network Traffic Forecasting using Adaptive Gradient Based Forecasting Model (AGBFM)

Kuldeep Singh Malik[1], Dr. Tilak Raj Rohilla[2]and Dr. Sandeep Kumar[3]

[1]Research Scholar, BabaMastnath University, Asthal Bohar, Rohtak

[2]Assistant Professor, Baba Mastnath University, Asthal Bohar, Rohtak

[3]Associate Professor, Maharaja Surajmal Institute of Technology, Janakpuri, New Delhi

ABSTRACT

The use of mobile phones is increasing rapidly with varying traffic demand. This exponential increase in mobile traffic is due to video streaming, mobile TV, video conferencing etc. 4G network is unable to full fill this increasing demand of bandwidth, broader coverage and lower latency. Due to this ever increasing demand, next generation of mobile networks which uses SDN, NFV, and cloud approach in mobile networks came into picture. Forecasting of forthcoming network traffic conditions is of great significance for improving performance of current networks. Efficient traffic forecasting technique becomes a challenging task for 5G networks. In this regard, adaptive optimization based training models for forecasting is proposed. In this paper, Adaptive Gradient Based Forecasting Model (AGBFM) is proposed and implemented. The paper also presents the results obtained after implementation of the proposed model, in terms of confidence interval and safe zone.

KEYWORDS*: Software Defined Network, 5G Networks, Gradient Network Optimizer*

INTRODUCTION

This paper presents the proposed Adaptive Gradient Based Forecasting Model (AGBFM) for network traffic forecasting. The proposed gradient based model is implemented. The model is designed and implemented using the two scenarios of datasets as explained below.

Dataset Generation Case 1: Raw data is generated by flooding various virtual machines with network traffic. The network traffic could be generated by running various commands at command prompt of the host machine. Here "ping command" and "apache server" 50 commands are run again and again to flood network with traffic. These bursts of data are captured by the host machine by using Wireshark (Soepeno, 2023). The Wireshark capture engine captures live network data simultaneously from multiple network interfaces.

Dataset Generation Case 2: The data used to evaluate the present study is taken from CRAWDAD (Community Resource for Archiving Wireless Data) (Sengupta et al., 2015)iitkdg/apptraffic datasets of a Smartphone app collected using tcpdump. Tcpdump is command-line interface-based data-network packet analyzer computer software. It allows the user to see TCP/IP and other packets that are being sent and received across a network to which the machine is connected. The desired traffic for evaluation came from Google Hangout of Smartphone app. An application called Google hangout (Bolton, 2014) traffic (GB/micro-sec) time(micro-seconds) traffic (GB/micro-sec) time(micro-seconds) 52 facilitates its users to do chats, carry out VoIP calls and video calls. Google hangout is not a completely peer-to-peer service platform, although it has features of a peer-topeer application as it permits two users to interconnect in real-time using a session server which is selected dynamically. This is the reason Google hangout application is called semi peer-to-peer platform. In the form of .pcap files, data was collected with only the Google Hangouts app running in the foreground and only required system functions running in the background. A subset of dataset form Google hangouts of Smartphone app collected from CRAWDAD community.

This proposed model is based on training using least square support vector machine (LSSVM). Further, a newly proposed adaptive gradient based optimizer is hybridized with LSSVM to enhance the performance of forecasting in case of network traffic. This forecasting model can be implemented at the application layer as explained in figure 1.

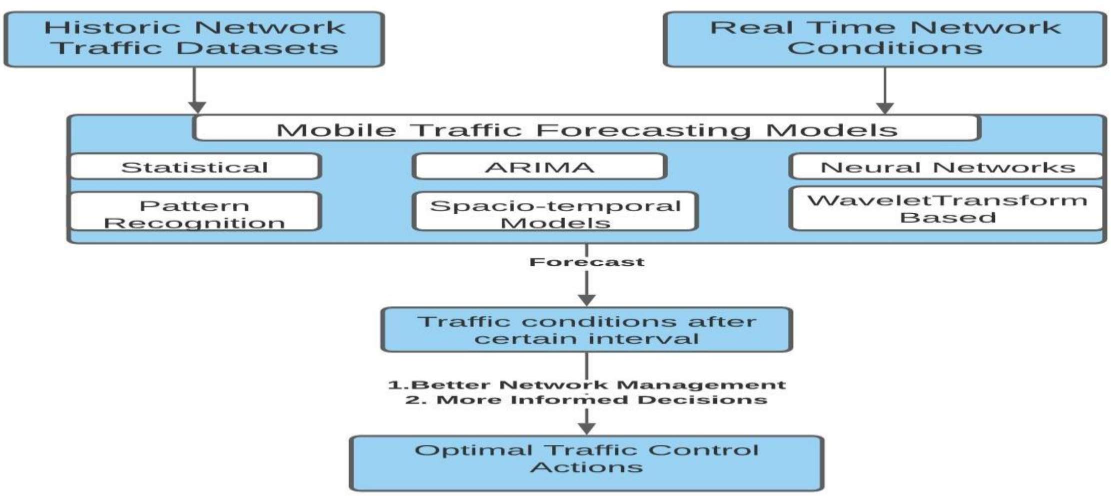


Fig 1. Architectural Flow of the Proposed Model

This paper also presents the results obtained after implementation of the proposed model, in terms of confidence interval and convergence curve. Further, the testing results using testing data in terms of confusion matrix, accuracy, prediction, sensitivity, and f1-score for the proposed

AGBFM model are also presented. The paper is further organized as follows. Section 2 shows the related work. Section 3 describes the forecasting model. In Sect. 4, proposed AGBFM is described followed by simulation results. Section 5 concludes the results.

RELATED WORK

The most popular traffic forecasting technique involves non-heuristic optimization techniques like the Newton based techniques (Laverde, 2023), Quasi-Newton based techniques (Schittkowski &Zillober, 2005), the conjugate direction technique (R Hestenes & Stiefel, n.d.) and Levenberg Marquardt (LM) technique(Moré, 1978). These techniques, alone or in combination with AI have been practiced in many areas to determine results of different types of real world problems.

The Quasi-Newton technique is combined with PSO algorithm by Salajegheh and Salajegheh to improve the efficiency and accuracy of PSO algorithm (Salajegheh&Salajegheh, 2019). Ibtissem and Nouredine (Ibtissem&Nouredine, 2013) hybridized the Differential Evolution (DE) method with the gradient technique to enhance the exploitation capability in the native DE. Authors also employed an adaptive optimization model using the conjugate gradient(Shahidi et al., 2005. Researchers also developed a hybrid DE using an exploitation method called the conjugate gradient method in order to enhance the convergence power. These researches illustrated the importance of gradient-search methods.

EXISTING WORK ON GRADIENT BASED TECHNIQUES

The gradient-search techniques have been extensively developed to overcome optimization tasks. To find an optimal result using the gradient methods, a terminal point where the value of gradient comes out to be zero, must be determined. The conjugate method and Newton method of gradient based techniques were using the same concept. Gradient-search based method finds the direction of search and the process of search progresses in the selected direction for the optimal value. Analyzing the exploration information in gradient based approaches required the derivatives of the objective function along with the constraints.

Performance evaluation of five different stochastic gradient search optimization methods were studied by authors using convergence interval, total fluctuations, parameter change rate, number of iterations and particular test function values on a Stacked denoising autoencoder (SdA) model(Vincent et al., 2010). On the bases of experimental results, AdaDelta (Zeiler, 2012) showed a higher performance in terms of fast convergence as compared to the other optimizers. Though, the details of dataset used have not been shared. Authors performed a performance evaluation of four gradient based algorithms, specifically, native gradient, stochastic gradient, semi-stochastic gradient and stochastic average methods using softmax regression on MNIST digits' dataset and synthetic datasets (Liberti & Kucherenko, 2005). Stochastic gradient based method usually performs superior in the two tests carried out by the authors; however other two hybrid methods resulted in better accuracy within limited time. A relative performance evaluation of gradient descent and stochastic gradient decent was experimented in(Konečný &Richtárik, 2017) using accuracy and convergence time metric for performance measurement of linear and logistic regression on MNIST dataset. Authors also emphasized the challenges related to these optimizers along with approaches that may be employed to increase the optimizing power of gradient descent (Mustapha et al., 2020). Using real-world investigation, authors established sources of failures of gradient descent performance in deep learning in (Failures of Gradient-based Deep Learning, 2017).

From this existing literature, it appears that gradient based method may be explored extensively to obtain optimal hyper-parameters of the LSSVM training model so as to enhance the performance of the forecasting model in network traffic. Further the two main limitations of pure gradient based optimization as seen in the related work are: (1) slower convergence speed (2) tends to converge to the local optima. Thus, in the present study, the optimization using the gradient search methods is combined with population based optimizers to overcome the drawbacks of gradient search approaches for developing a prevailing and robust system. Therefore, it will be worthwhile to design optimizer which employ gradient search technique to avoid the unrealistic locations and tends toward the direction of realistic region. It also considers proficiencies of the population oriented approaches of optimization.

AGBFM forecasting model

Researchers have developed efficient mechanisms for the detection of outbreaks in traffic and subsequent traffic classification by using various machine learning models such as SVM and LSSVM. These forecasting models reduced the false alarms to a certain extent. Authors have also constructed various population optimization based models and redefined its computation rules to solve the Multi Objective Computational Problem. These models have definitely improved performance to an extent but these existing models have some limitations. The major limitations of the existing models are that many of these models are based primarily on some heuristic rules and do not actually rely on training the model based on past datasets and then predicting the future based on smart choice by modifying the model automatically as per the need of an hour. Also, these models have not considered any peaks in the traffic bursts. Therefore, a more traffic aware adaptive model is proposed, which tries to accommodate the traffic peaks and manage resources accordingly. Another challenge mobile networks are facing is, how to efficiently handle data oriented mobile communications, i.e. how to provide reasonable data performance to customers in networks when the cell sites experience significant amount of data traffic, resulting in call drops. Addressing these challenges requires a concrete understanding of the behavior of traffic networks and accurate mobile traffic forecast for efficient network planning and operations. The short term prediction of network traffic guides several immediate traffic bursts in the network. This understanding is used along with proposed adaptive gradient based forecasting model (AGBFM) in order to pre-process datasets for making network traffic frames to make it fit for training.

FLOW OF PROPOSED MODEL

The existing study presents various challenges and problems in identifying traffic patterns during seasonal events in networks. Therefore, a methodology is proposed so as to identify both short and long term traffic fluctuations in SDMN based architecture and to enable SDMN controller to forecast accurately. The brief view of complete methodology is shown in figure 2.

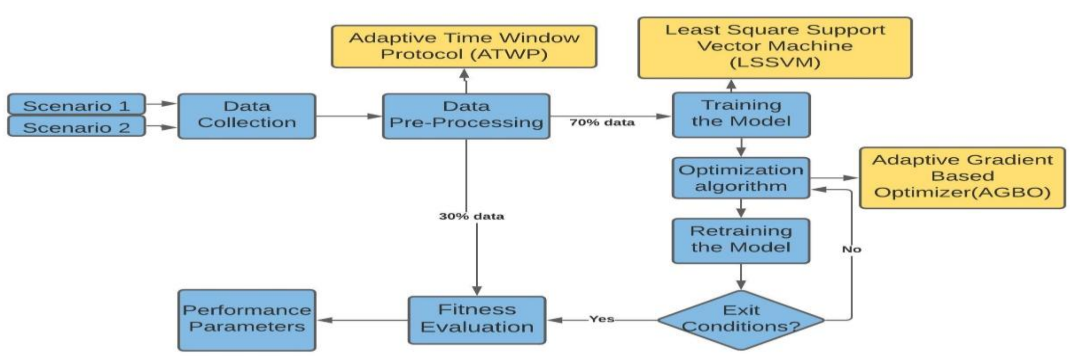


Fig 2. Brief view of proposed AGBFM model

After identification of classes of traffic, traffic bursts generation and obtaining initial value of parameters, prediction is done by traffic classes using the initial value of the tuning parameters. The proposed AGBFM uses Adaptive Gradient Based Optimization (AGBO) technique for optimizing the tuning parameters along with LSSVM for training the model. The flowchart in figure 3 describes the individual components of proposed methodology. It shows the flow after pre-processed data is obtained.

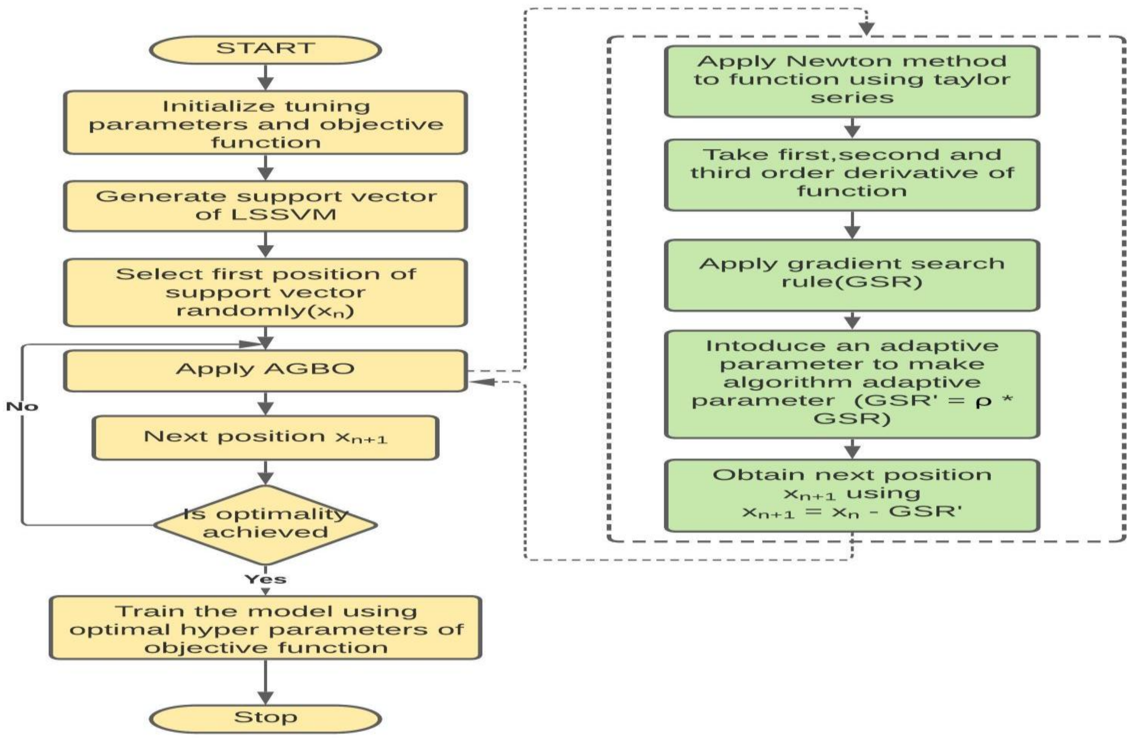


Fig 3. Detailed flow of AGBFM

MAJOR REQUIREMENTS OF PROPOSED MODEL

Major requirements of any forecasting model are:

- **i.** Data traffic for training and testing
- **ii.** Pre-processing procedure
- **iii.** Selecting machine learning training model
- **iv.** Selecting optimization technique for optimizing parameters of training
- **v.** Selecting exit conditions
- **vi.** Selecting performance parameters

Algorithm 1 shows the pseudo code of AGBFM network forecasting model.

Algorithm 1: Pseudo code for AGBFM

Parameters: ϑ (number of Support Vectors), *r*(traffic feature vectors)

Input: traffic dataset with η data points

Output: Class assignment for each data point *i* in dataset $\mathbb{C} \in$ Classes identified by ATWP)

Step 1. *Partition the dataset into training set X and test set X ˆ*

Step 2. *Identify mean μ and standard deviation δ in the dataset*

Step 3. *Normalize the dataset using min-max normalization*

Step 4. *Using Algorithm Time window Protocol (ATWP) estimate class ranges*

Step 5. *Initialize regression parameters γ and σ*

Step 6. For k = 1 to K (K-Fold validations)

a. ***e*** *= train using LSSVM on the pre-processed traffic dataset*

b. ***If*** *change in error e has not met the stopping conditions of optimization technique, then*

Re-initiate training parameter $\boldsymbol{\gamma}$ *and* $\boldsymbol{\sigma}^2$*using proposed AGBO optimization technique.*

c. ***Else***

Optimal training model has been received

Step 7. Exit for

Step 8. End for

Step 9. Test AGBFM model on test dataset$\widehat{X}$

Step 10. Get estimated traffic ranges $\widehat{\mathcal{Y}} = AGBFM(\widehat{X})$, $\widehat{\mathcal{Y}}$ *consists of* $\mathbb{C}$*classes*

Step 11. Evaluate using true class $\mathcal{Y}$ *and predicted class* $\widehat{\mathcal{Y}}$*using various performance metrics*

RESULT AND ANALYSIS

The following are the simulation results of the proposed forecasting model AGBFM:

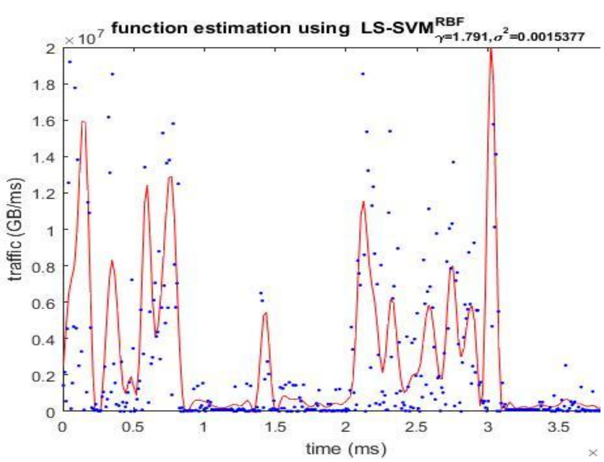
SIMULATION PARAMETERS

The two scenarios of datasets are considered to evaluate the performance of the proposed model over different parameters. For the performance evaluation, MATLAB 2016 on Intel i5 processor with 8GB of RAM is used in this study. The parameters used in the implementation are listed in table 1.

Parameter	Value
TWP Window	1e^{-4}
K-Cross validation folds for LSSVM	15
No of iterations	Depends on optimality
Cost Function	MSE and RMSE
Dataset size	Scenario 1: 78120 records Scenario 2: 56556 records
Training Model	Least Square Support Vector Machine (LSSVM)
Optimization function	Adaptive Gradient Based Optimizer (AGBO)

Table 1. Simulation parameters considered in the proposed work

Scenario 1: The predicted model using finally optimized hyper-parameters value i.e. regularization factor or punishment factor value as 1.791 and bandwidth of RBF kernel as 0.0015377 is shown in figure 4. The safe zone and confidence interval of the traffic in scenario 1 is shown in figure 5.



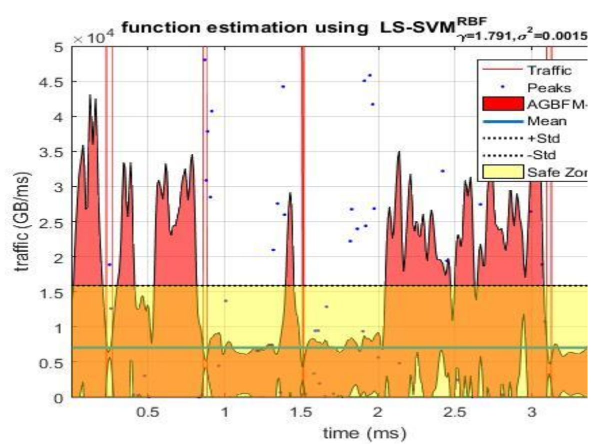


Figure 4. AGBFM model prediction for network traffic in scenario 1

Figure 5. **C**lassification of AGBFM model traffic prediction in scenario 1

Scenario 2: The predicted model using finally optimized hyper-parameters value i.e. regularization factor or punishment factor value as 0.83452 and bandwidth of RBF kernel as 0.00013626 is shown in figure 6. The safe zone and confidence interval of the traffic in scenario 2 is shown in figure 7.

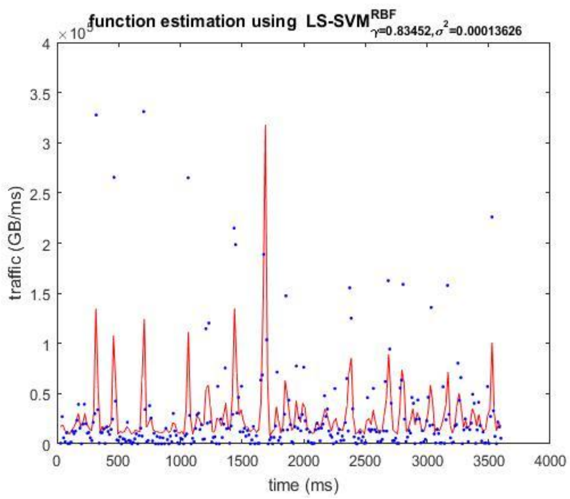


Fig 6. AGBFM model prediction for scenario scenario 2

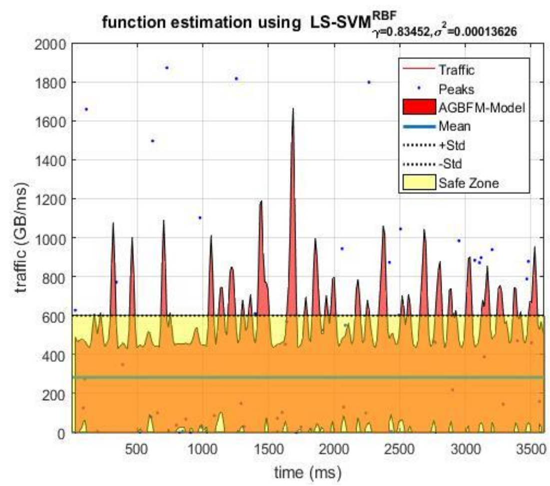


Fig 7.Classification of network traffic in AGBFM traffic prediction in scenario 2

CONCLUSION

The increasing demand of customers for mobile network services and applications has further emphasized the need of high speed networks. One of the way out to address this ever increasing demand is to forecast traffic and use the available resources efficiently and effectively. So the network technologists came up SDMN. SDMN technology implements various services at its application layer in the form of software. In this research, traffic flow forecasting models to be implemented at the application layer of SDMN controller are proposed. These models provide information about future network load and raise appropriate alarms so that appropriate decisions could be taken by the controller to redirect network traffic in case of congestion at a particular link.

The research work presented here discusses some important aspects related to efficient mobile traffic forecasting using Software Defined Mobile Network paradigm. A central theme of the paper is to design and implement an efficient forecasting model for today's networks that can take appropriate decisions about network incoming traffic. The main aim to develop the model is to minimize burden of explosively growing incoming network traffic on the available resources of a particular network and henceforth, provide the users with good services at the network. The existing traffic forecasting algorithms faces scalability issues while adding more resources and lesser agility to modify model. Also, in the existing forecasting algorithms the complexity in terms of cost, time and memory consumption is high. The self-similarity and periodicity of traffic data is a major issue in applying various algorithms. Periodicity is an indicator of malware command, network congestion and denial of service attacks.

Forecasting model AGBFM have been proposed and implemented. The proposed model considered LSSVM machine learning algorithm for training using datasets. Models are tuned with the value of initial regularization parameter (γ) and kernel width parameter (σ^2) also called width parameter.AGBFM model uses preprocessed data to train the model using LSSVM with Adaptive Gradient Based Optimizer to optimize the parameters of model. In future the work can be extended to encompass:

- ❖ Deployment of forecasting models: A well-planned deployment is required which constrains tighter interaction with the existing network. The application to be deployed should interact with existing software infrastructure and it should digest existing temporal data.
- ❖ Optimization parameters: Efficient techniques are required for optimizing more number of parameters of the training model used in forecasting model

REFERENCES

1. Bolton, R. (2014). Google Hangouts. *The Journal of the Canadian Health Libraries Association*, 34(1), 39. [Cited above 2024 April, 10]; available from https://doi.org/10.5596/c13-002.pdf

2. Failures of Gradient-based Deep Learning. (2017). [Cited above 2024 April, 10]; available from https://dl.acm.org/doi/abs/10.5555/3305890.3305998.pdf

3. Ibtissem, C., &Nouredine, L. (2013). A hybrid method based on conjugate gradient trained neural network and differential evolution for non linear systems identification. In

Conference: Electrical Engineering and Software Applications (ICEESA), 2013. https://doi.org/10.1109/iceesa.2013.6578397

4. Konečný, J., &Richtárik, P. (2017). Semi stochastic gradient descent methods. *Frontiers in Applied Mathematics and Statistics*, *3*. [Cited above 2024 April, 10]; available fromhttps://doi.org/10.3389/fams.2017.00009.pdf

5. Laverde, J. G. T. (2023). On the Newton-Raphson method and its modifications. *CIENCIA EN DESARROLLO*, *14*(2), 75–80. [Cited above 2024 April, 10]; available from https://doi.org/10.19053/01217488.v14.n2.2023.15157

6. Liberti, L., & Kucherenko, S. (2005). Comparison of deterministic and stochastic approaches to global optimization. *International Transactions in Operational Research*, *12*(3), 263–285. [Cited above 2024 April, 10]; available from https://doi.org/10.1111/j.1475-3995.2005.00503.x.pdf

7. Moré, J. J. (1978). The Levenberg-Marquardt algorithm: Implementation and theory. In *Lecture notes in mathematics* (pp. 105–116). [Cited above 2024 April, 10]; available from https://doi.org/10.1007/bfb0067700

8. Mustapha, A., Mohamed, L., & Ali, K. (2020). An overview of gradient descent algorithm optimization in Machine Learning: Application in the ophthalmology field. In *Communications in computer and information science* (pp. 349–359). [Cited above 2024 April, 10]; available from https://doi.org/10.1007/978-3-030-45183-7_27.pdf

9. RHestenes, M., & Stiefel, E. (n.d.). Methods of Conjugate Gradients for Solving Linear Systems 1. *Journal of Research of the National Bureau of Standards*, *Vol. 49,*(No.6), 409–436. [Cited above 2024 April, 10]; available from https://nvlpubs.nist.gov/nistpubs/jres/049/jresv49n6p409_a1b.pdf

10. Salajegheh, F., &Salajegheh, E. (2019). PSOG: Enhanced particle swarm optimization by a unit vector of first and second order gradient directions. Swarm and *Evolutionary Computation*, *46*, 28–51. [Cited above 2024 April, 10]; available from https://doi.org/10.1016/j.swevo.2019.01.010

11. Schittkowski, K., &Zillober, C. (2005). Nonlinear programming: algorithms, software, and applications. In *IFIP International Federation for Information Processing/IFIP* (pp. 73–107). [Cited above 2024 April, 10]; available from https://doi.org/10.1007/0-387-23467-5_5

12. Sengupta, S., Gupta, H., Ganguly, N., Mitra, B., De, P., & Chakraborty, S. (2015). CRAWDAD dataset iitkgp/apptraffic (v.2015-11-26). *CRAWDAD Wireless Network Data Archive*. [Cited above 2024 April, 10]; available from https://doi.org/10.15783/c77s3w.pdf

13. Shahidi, N., Esmaeilzadeh, H., Abdollahi, M., Ebrahimi, E., & Lucas, C. (2005). Self-adaptive memetic algorithm: an adaptive conjugate gradient approach. In *IEEE Conference on Cybernetics and Intelligent Systems* (Vol. 2). [Cited above 2024 April, 10]; available from https://doi.org/10.1109/iccis.2004.1460378.pdf

14. Soepeno, R. (2023). Wireshark: an effective tool for network analysis. *ResearchGate*. [Cited above 2024 April, 10]; available from https://doi.org/10.13140/RG.2.2.34444.69769.pdf

15. Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., &Manzagol, P.-A. (2010). Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. *Journal of Machine Learning Research*, *11*. [Cited above 2024 March, 10]; available from https://www.jmlr.org/papers/volume11/vincent10a/vincent10a.pdf

16. Zeiler, M. D. (2012). ADADELTA: an Adaptive Learning Rate Method. *(Cornell University)*. [Cited above 2024 March, 10]; available from https://doi.org/10.48550/arxiv.1212.5701.pdf