Performance Analysis of Scaled Conjugate Gradient (SCG) Algorithm on Computing Problems

A A G B Ariana[1,a)], Irzal Arief Wisky[2,b)], N L W S R Ginantra[3,c)], Moch. Rachmandany Firmansyah[4,d)], Achmad Daengs GS[5,e)]

[1] STMIK STIKOM Indonesia, Bali, Indonesia
[2] Universitas Putra Indonesia YPTK, Padang, Indonesia
[3]STIKI Indonesia, Bali, Indonesia
[4] Universitas Airlangga, Surabaya.Indonesia
[5] Universitas 45 Surabaya, Surabaya, Indonesia

[a)] Corresponding author: gungariana@stiki-indonesia.ac.id
[b)] irzal.arief12@gmail.com
[c)] wiwik@stiki-indonesia.ac.id
[d)] bumigora80@gmail.com
[e)] adaengsgs@univ45sby.ac.id

Abstract. The artificial neural network has several training functions that can speed up the training process of the standard backpropagation algorithm. Therefore, the purpose of this research is to evaluate the scaled conjugate gradient algorithm's capability and performance, which develops the training function of standard backpropagation to solve computational problems. The dataset used in this paper uses quantitative data from export data of jewelry and valuable goods by the leading destination country, which is processed from documents customs of the Directorate General of Customs and Excise (PEB and PIB) and quoted from the Indonesian Statistical Publication. A network architecture model will be formed and determined based on this data, including 7-7-1, 7-14-1, and 7-21-1. Based on these three models after training and testing, the results show that the model with the best performance and accuracy is 7-14-1 with a performance value of 0.001118426, the lowest among the three other models and an accuracy of 90.9% (higher than the two models other). So it can be concluded that the SCG algorithm with the 7-14-1 model can be used to solve computational problems, as evidenced by the best performance and accuracy values.

Keywords: Performance, SCG, Computing, Forecasting, Machine Learning

INTRODUCTION

Computing is described as the process of deriving solutions to issues from input data via the use of algorithms [1]. This is called computational theory, a subfield of mathematics and computer science. Thousands of years ago, calculations and computations were mainly performed using pen and paper, chalk and slate, or mentally, occasionally with the assistance of a table. However, computers have been used to perform a large amount of computation in the modern era [2]. In practical use, it usually applies computer simulations or various other forms of computing to solve problems in various scientific fields. Still, it is also used to find fundamental new principles in science in its development. This field is different from computer science, which studies computing, computers, and data processing. Additionally, this discipline is distinct from theory and experiment, which are considered classic types of science and scientific endeavor [3]. In the natural sciences, computational science can bring new insights by utilizing mathematical models embedded in computer programs that are built on established theoretical foundations to tackle real-world issues [4]. The computation proposed in this paper uses ANN or Machine learning Algorithm SCG. The discussion is devoted to looking at the ability or performance of the algorithm to compute data in determining the best model that can be used for forecasting problems. SCG has an advantage over other conjugate gradient algorithms (Fletcher-Reeves, Polak-Ribiere, and Powell-Beale), namely that each iteration does not search by row. Based on the row search, all training inputs from the network response are calculated several times for each lookup, thus requiring a long computational time. So, to avoid long queues that take up searching time, the SCG algorithm is used. SCG is fully automated and is a supervised algorithm. This means that critical user-dependent parameters are not included, which is much faster than Levenberg Backpropagation. This algorithm can be used on any dataset, if the input, transfer function, and weights of the given dataset have derivative functions [5–9].

Several related studies are the background of this research, including The use of the SCG algorithm in the network to train sample features quickly without other parameters such as learning speed for Rotating Machine Fault Diagnosis. This study uses three methods: the segmentation method to increase extraction, the augmentation method to increase the number of samples, and SCG to accelerate learning. The result is that the way obtains superior performance for feature learning and classification [10]. The following study modified the SCG method with a non-convex objective function for large-scale optimization without constraints. Based on the proposed method emerges an extraordinary feature that has global convergence on the aim function even without convexity assumption. The numerical results show that the new modification based on the proposed research of SCG is efficient [11]. The following research is to detect and diagnose breast cancer using SCG and Advanced Deep Learning Algorithm. How accurate is the method for detecting breast cancer? The proposed layered neural network in the form of a model uses this algorithm to perform an optimal classification of cancer cells by considering several appropriate parameters. Findings: The classified cancer cells were then evaluated using the proposed algorithm with the designed layered neural network model. After the model training model is successful, the model classifies cancer as benign (2) or malignant (4). Based on the proposed method, eleven input parameters have been taken for the better calculation. classification of artificial neural network patterns based on performance analysis of the proposed SCG model optimization technique was carried out by taking 699 data and classified as benign [12].

Based on related descriptions from previous studies, this article will examine the SCG algorithm's performance in solving computational issues, which will later produce predictive models, especially for export data of jewelry and valuable goods by the leading destination country. This dataset is only used to aid in the verification and process of measuring the algorithm's performance. The purpose of this study is to determine the accuracy and implementation of the algorithm in finding the best results to solve forecasting (prediction) problems.

RESEARCH METHODS

Research Materials

The dataset used in this paper uses quantitative data, in the form of export data of jewelry and valuable goods by the leading destination country, which is processed from documents customs of the Directorate General of Customs and Excise (PEB and PIB) and quoted from the Indonesian Statistical Publication, which can be seen in table 1 [13].

TABLE 1. Exports of Jewelry and Valuables by Main Destination Countries, 2012-2020 (Tons)

No	Country of Destination	2012	2013	2014	2015	2016	2017	2018	2019	2020
1	Switzerland	1.9	3.2	3.3	26.3	54.1	29.0	13.5	5.6	5.9
2	Singapore	110.3	124.2	129.8	136.4	120.4	102.7	139.7	105.6	119.4
3	Hong Kong	8.6	10.6	47.1	76.8	83.8	59.5	87.6	79.5	16.3
4	UEA	2.5	0.8	12.1	8.9	8.1	9.8	6.3	5.1	2.2
5	South Africa	4.0	9.0	30.6	11.6	12.6	4.0	6.5	2.7	2.7
6	Taiwan	1.4	0.5	10.1	24.5	3.9	1.1	0.2	0.1	1.5
7	USA	142.7	153.0	187.8	226.0	209.6	157.7	244.1	215.1	172.7
8	India	18.7	0.9	1.6	16.9	8.6	3.3	46.6	0.4	0.1
9	Australia	18.3	29.2	42.5	30.5	33.2	30.6	28.6	17.2	20.2
10	Italy	18.9	26.7	30.6	32.3	36.6	35.6	24.9	14.2	12.2
11	Other	4266.7	440.9	438.5	378.8	644.4	488.1	435.4	379.5	261.8

Research Stages

The stages of the research were carried out to achieve the research objectives which can be seen in figure 1.

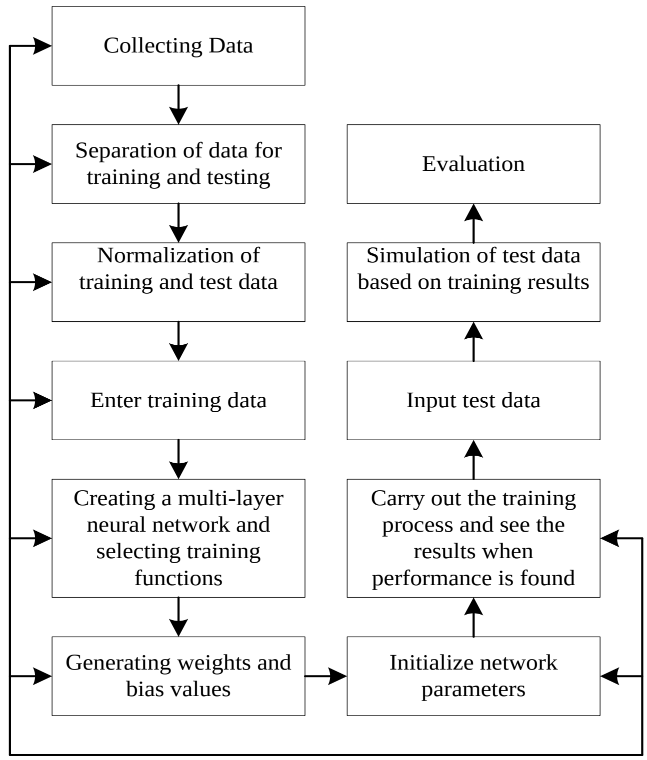

FIGURE 1. General Research Method Flowchart

Based on Figure 1, it can be explained that the first step taken from the research stage is to collect the research dataset (table 1). The next step is to separate the research dataset into two groups: training and testing data. The next stage is to normalize the training and testing data using the equation formula (1) [14–21].

$$x' = \frac{0.8\,(x-a)}{b-a} + 0.1 \quad (1)$$

Where : x' is the result of data that has been normalized, 0.8 and 0.1 are the default values of the normalization formula, x is the data to be normalized, a is the lowest value of the dataset, and b is the highest value of the dataset. Furthermore, the normalized training data is entered into the Matlab 2011b application for processing, creating a multi-layer neural network (training data input). Next is the application of the SCG algorithm. The creation of this multi-layer neural network uses the tansig and logsig functions. Next is to initialize the network parameters based on the training function (trainscg). Then enter the command to carry out the training process and see the results when performance is found. If the training results reach convergence, it will continue to enter the normalized test data. But if the training results have not reached convergence, return to the initialization stage of network parameters. The next stage is followed by a simulation of test data based on the training results. If everything has been done, the final step is to evaluate the best architectural model based on the lowest (small) Performance/MSE test and based on the best accuracy.

RESULTS AND DISCUSSION

Normalization and Data Processing

The study dataset in Table 1 must be normalized using the following equation (1) so that a value is obtained between 0 to 1. The training data is based on data for 2012-2018 (N1-X7) with a training target of 2019 (A1), while the test data is based on 2013-2019 data (N8-N14) and test target of 2020 (A2). The data analysis process using Microsoft Excel and Matlab 2011b, with the SCG algorithm. The network architecture that will be used for the training process in this research is 7-7-1, 7-14-1, and 7-21-1 (7 is the input layer, 7, 14, and 21 is the hidden layer with 7, 14, and 21 neurons, 1 is the output layer).

Based on the analysis of training and testing that has been done, the best network architecture model is 7-14-1. The results of training and testing can be seen in table 2.

TABLE 2. Training Data

No	N1	N2	N3	N4	N5	N6	N7	Target (A1)	Epoch 138		
									Real	Mistake	SSE
1	0,1003	0,1006	0,1006	0,1049	0,1101	0,1054	0,1025	0,1010	0,1027	-0,0017	0,000002785
2	0,1207	0,1233	0,1243	0,1256	0,1226	0,1192	0,1262	0,1198	0,1196	0,0002	0,000000033
3	0,1016	0,1020	0,1088	0,1144	0,1157	0,1111	0,1164	0,1149	0,1143	0,0006	0,000000345
4	0,1005	0,1001	0,1023	0,1017	0,1015	0,1018	0,1012	0,1009	0,0987	0,0022	0,000005006
5	0,1007	0,1017	0,1057	0,1022	0,1023	0,1007	0,1012	0,1005	0,1005	0,0000	0,000000000
6	0,1002	0,1001	0,1019	0,1046	0,1007	0,1002	0,1000	0,1000	0,1000	0,0000	0,000000000
7	0,1267	0,1287	0,1352	0,1424	0,1393	0,1296	0,1458	0,1403	0,1404	-0,0001	0,000000008
8	0,1035	0,1002	0,1003	0,1032	0,1016	0,1006	0,1087	0,1001	0,1004	-0,0003	0,000000118
9	0,1034	0,1055	0,1080	0,1057	0,1062	0,1057	0,1053	0,1032	0,1036	-0,0004	0,000000155
10	0,1035	0,1050	0,1057	0,1060	0,1068	0,1067	0,1047	0,1026	0,1032	-0,0006	0,000000309
11	0,9000	0,1827	0,1822	0,1710	0,2208	0,1915	0,1816	0,1711	0,1711	0,0000	0,000000001
									Sum SSE		0,000008761
									Perf/ MSE		0,000000796

TABLE 3. Testing Data

X	N8	N9	N10	N11	N12	N13	N14	Target (A2)	Real	Mistake	SSE	Hasil
1	0,1006	0,1006	0,1049	0,1101	0,1054	0,1025	0,1010	0,1011	0,1057	-0,0046	0,000021275	1
2	0,1233	0,1243	0,1256	0,1226	0,1192	0,1262	0,1198	0,1224	0,1134	0,0090	0,000080445	1
3	0,1020	0,1088	0,1144	0,1157	0,1111	0,1164	0,1149	0,1030	0,1110	-0,0080	0,000063401	1
4	0,1001	0,1023	0,1017	0,1015	0,1018	0,1012	0,1009	0,1004	0,0995	0,0009	0,000000799	1
5	0,1017	0,1057	0,1022	0,1023	0,1007	0,1012	0,1005	0,1005	0,1010	-0,0005	0,000000263	1
6	0,1001	0,1019	0,1046	0,1007	0,1002	0,1000	0,1000	0,1003	0,0992	0,0011	0,000001129	1
7	0,1287	0,1352	0,1424	0,1393	0,1296	0,1458	0,1403	0,1324	0,1395	-0,0071	0,000050937	1
8	0,1002	0,1003	0,1032	0,1016	0,1006	0,1087	0,1001	0,1000	0,0970	0,0030	0,000009000	1
9	0,1055	0,1080	0,1057	0,1062	0,1057	0,1053	0,1032	0,1038	0,1038	0,0000	0,000000001	1
10	0,1050	0,1057	0,1060	0,1068	0,1067	0,1047	0,1026	0,1023	0,1043	-0,0020	0,000004126	1
11	0,1827	0,1822	0,1710	0,2208	0,1915	0,1816	0,1711	0,1491	0,0392	0,1099	0,012071311	0
									Sum SSE		0,012302685	90,9%
									Perf / MSE		0,001118426	

Table 2 and Table 3 are the best results from the SCG algorithm which is entered into the Matlab application which is then checked again for validity using Ms. Excel. The actual value is obtained from the output of training and testing results. The Mistake value is obtained from the calculation of Target (A) - Actual. The SSE value is obtained based on the calculation of Mistake^ 2. The number of SSE is obtained from the sum of the overall SSE values. Meanwhile, the Performance/MSE value is obtained from the Total SSE/11 (11 is the number of data records). The results are obtained based on the reference from the mistake value, if the Mistake value is <= 0.01 then it produces a value of 1, if not then it produces a value of 0. Accuracy of 90.9% is obtained from the number of results/11*100. Based on the results of training and testing between the Matlab application and Ms Excel, it turned out to be valid.

Comparison of Algorithm Results

The comparison of the results of the SCG algorithm using the 7-7-1, 7-14-1, and 7-21-1 network architecture models can be seen in table 4.

TABLE 4. Comparison of SCG Algorithm Results

Algorithm	Model	Training Function	Transfer Function	Epoch (Iteration)	MSE Training	MSE Testing / Performance	Accuracy
Scaled	7-7-1	trainscg	tansig, logsig	157	0,000000220	0,001437091	63,6%
Conjugate	7-14-1	trainscg	tansig, logsig	138	0,000000796	0,001118426	90,9%

Algorithm	Model	Training Function	Transfer Function	Epoch (Iteration)	MSE Training	MSE Testing / Performance	Accuracy
Gradient	7-21-1	trainscg	tansig, logsig	386	0,000000015	0,003109785	72,7%

While the comparison graph using the 7-7-1, 7-14-1, and 7-21-1 network architecture models can be seen in Figure 2.

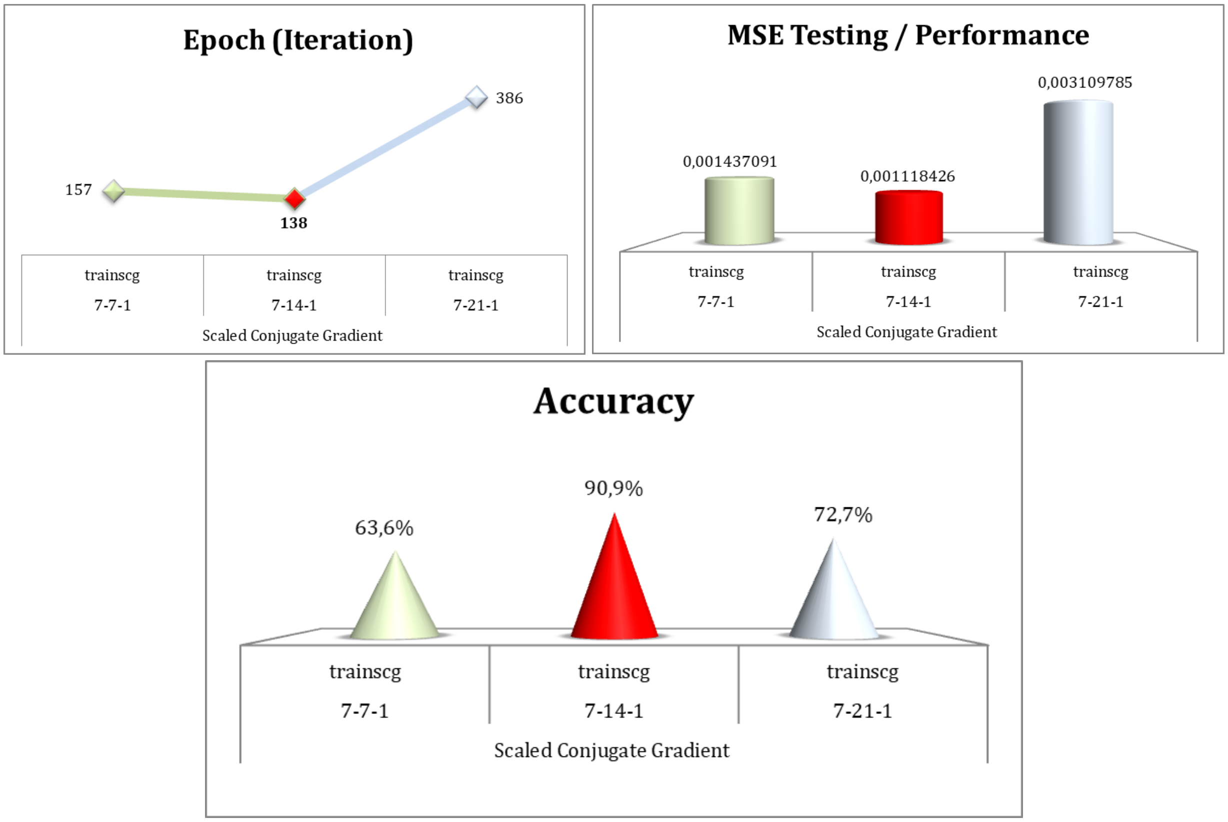

FIGURE 2. Comparison Chart

Based on table 4 and figure 2, the comparison of each model used with the SCG algorithm (trainscg) can be seen. 7-14-1 architectural model is the best compared to other architectural models. The resulting epoch is 138 iterations, the value of MSE Testing (Performance) is 0.001118426, and the highest accuracy compared to other models (90.9%). When viewed from the MSE Training, the 7-21-1 architectural model is better, but it is not stable when tested. Another reason that makes the 7-14-1 architectural model chosen as the best model is the consistency and stability of the computations during the training and testing process.

CONCLUSIONS

It can be concluded in accordance with the aims and objectives of the study, that the SCG Algorithm can be used to solve computational problems because the training time to achieve convergence is not too long and the resulting performance is quite good. This is based on the analysis that the SCG Algorithm (trainscg) produces not too many iterations, a low Performance / MSE test value and a high level of accuracy.

REFERENCES

1. X. Xu *et al.*, "A computation offloading method over big data for IoT-enabled cloud-edge computing," *Future Generation Computer Systems*, vol. 95, pp. 522–533, 2019.
2. S. K. Sen and R. P. Agarwal, *Computing: Birth, growth, exaflops computation and beyond*, vol. 247.

2020.
3. L. Grozinger *et al.*, “Pathways to cellular supremacy in biocomputing,” *Nature Communications*, vol. 10, no. 1, pp. 1–11, 2019.
4. A. Johanson and W. Hasselbring, “Software Engineering for Computational Science: Past, Present, Future,” *Computing in Science and Engineering*, vol. 20, no. 2, pp. 90–109, 2018.
5. G. Yuan, T. Li, and W. Hu, “A conjugate gradient algorithm for large-scale nonlinear equations and image restoration problems,” *Applied Numerical Mathematics*, vol. 147, no. 11661009, pp. 129–141, 2020.
6. D. Selvamuthu, V. Kumar, and A. Mishra, “Indian stock market prediction using artificial neural networks on tick data,” *Financial Innovation*, vol. 5, no. 1–16, pp. 1–12, 2019.
7. C. B. Khadse, M. A. Chaudhari, and V. B. Borghate, “Electromagnetic Compatibility Estimator Using Scaled Conjugate Gradient Backpropagation Based Artificial Neural Network,” *IEEE Transactions on Industrial Informatics*, vol. 13, no. 3, pp. 1036–1045, 2017.
8. L. Babani, S. Jadhav, and B. Chaudhari, “Scaled Conjugate Gradient Based Adaptive ANN Control for SVM-DTC Induction Motor Drive,” *IFIP Advances in Information and Communication Technology*, vol. 475, pp. 384–395, 2016.
9. Q. Huang and L. Cui, “Design and application of face recognition algorithm based on improved backpropagation neural network,” *Revue d’Intelligence Artificielle*, vol. 33, no. 1, pp. 25–32, 2019.
10. Z. Meng, X. Guo, Z. Pan, D. Sun, and S. Liu, “Data Segmentation and Augmentation Methods Based on Raw Data Using Deep Neural Networks Approach for Rotating Machinery Fault Diagnosis,” *IEEE Access*, vol. 7, pp. 79510–79522, 2019.
11. Z. Khoshgam and A. Ashrafi, “A new modified scaled conjugate gradient method for large-scale unconstrained optimization with non-convex objective function,” *Optimization Methods and Software*, vol. 34, no. 4, pp. 783–796, 2019.
12. P. K. Vadla, Y. V. R. N. Pawan, B. P. Kolla, and S. L. Tripathi, “Accurate Detection and Diagnosis of Breast Cancer Using Scaled Conjugate Gradient Back Propagation Algorithm and Advanced Deep Learning Techniques,” in *Advances in Electrical and Computer Technologies. Lecture Notes in Electrical Engineering*, vol. 711, 2021, pp. 99–112.
13. Badan Pusat Statistik, “Ekspor Barang Perhiasan dan Barang Berharga Menurut Negara Tujuan Utama, 2012-2020,” *Publikasi Statistik Indonesia*, 2020. [Online]. Available: https://www.bps.go.id/statictable/2019/02/25/2028/ekspor-barang-perhiasan-dan-barang-berharga-menurut-negara-tujuan-utama-2012-2020.html. [Accessed: 13-Dec-2021].
14. G. W. Bhawika *et al.*, “Implementation of ANN for Predicting the Percentage of Illiteracy in Indonesia by Age Group,” *Journal of Physics: Conference Series*, vol. 1255, no. 1, pp. 1–6, 2019.
15. A. Wanto *et al.*, “Analysis of the Backpropagation Algorithm in Viewing Import Value Development Levels Based on Main Country of Origin,” *Journal of Physics: Conference Series*, vol. 1255, no. 1, pp. 1–6, 2019.
16. E. Siregar, H. Mawengkang, E. B. Nababan, and A. Wanto, “Analysis of Backpropagation Method with Sigmoid Bipolar and Linear Function in Prediction of Population Growth,” *Journal of Physics: Conference Series*, vol. 1255, no. 1, pp. 1–6, 2019.
17. M. K. Z. Sormin, P. Sihombing, A. Amalia, A. Wanto, D. Hartama, and D. M. Chan, “Predictions of World Population Life Expectancy Using Cyclical Order Weight / Bias,” *Journal of Physics: Conference Series*, vol. 1255, no. 1, pp. 1–6, 2019.
18. A. Wanto *et al.*, “Analysis of the Accuracy Batch Training Method in Viewing Indonesian Fisheries Cultivation Company Development,” *Journal of Physics: Conference Series*, vol. 1255, no. 1, pp. 1–6, 2019.
19. Y. Andriani, H. Silitonga, and A. Wanto, “Analisis Jaringan Syaraf Tiruan untuk prediksi volume ekspor dan impor migas di Indonesia,” *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, vol. 4, no. 1, pp. 30–40, 2018.
20. W. Saputra, J. T. Hardinata, and A. Wanto, “Implementation of Resilient Methods to Predict Open Unemployment in Indonesia According to Higher Education Completed,” *JITE (Journal of Informatics and Telecommunication Engineering)*, vol. 3, no. 1, pp. 163–174, 2019.
21. N. L. W. S. R. Ginantra *et al.*, “Performance One-step secant Training Method for Forecasting Cases,” *Journal of Physics: Conference Series*, vol. 1933, no. 1, pp. 1–8, 2021.