Article

Introducing an evolutionary method to create the bounds of artificial neural networks

Ioannis G. Tsoulos[1,*], **Vasileios Charilogis**[2], **Dimitrios Tsalikakis**[3]

[1] Department of Informatics and Telecommunications, University of Ioannina, Greece; itsoulos@uoi.gr
[2] Department of Informatics and Telecommunications, University of Ioannina, Greece; v.charilog@uoi.gr
[3] Department of Engineering Informatics and Telecommunications, University of Western Macedonia, 50100 Kozani, Greece; tsalikakis@gmail.com
* Correspondence: itsoulos@uoi.gr

Abstract: Artificial neural networks are widely used in applications from various scientific fields and in a multitude of practical applications. In recent years, a multitude of scientific publications have been presented on the effective training of their parameters, but in many cases overfitting problems appear, where the artificial neural network shows poor results when used on data that was not present during training. This text proposes the incorporation of a three - stage evolutionary technique, which has its roots in the differential evolution technique, for the effective training of the parameters of artificial neural networks and the avoidance of the problem of overfitting. The new method effectively constructs the parameter value range of the artificial neural network, achieving both a reduction in training error and preventing the network from experiencing overfitting phenomena. This new technique was successfully applied to a wide range of problems from the relevant literature and the results were extremely promising.

Keywords: Neural networks; Evolutionary algorithms; Stochastic methods; Differential Evolution

Citation: Tsoulos, I.G.; Charilogis, V.; Tsalikakis D. Introducing an evolutionary method to create the bounds of artificial neural networks. *Journal Not Specified* **2024**, *1*, 0. https://doi.org/

Received:
Revised:
Accepted:
Published:

1. Introduction

The remaining of this article is organized as follows: in section 2 the proposed method is discussed in detail, in section 3 the used datasets as well as the conducted experiments are discussed and finally, in section 4 some conclusions are presented.

2. Materials and Methods

3. Results

The new training procedure for the construction of weights of neural networks was tested on a wide series of classification and regression datasets, found in the recent literature and in the following websites:

1. The UCI website, https://archive.ics.uci.edu/(accessed on 24 November 2024)[1]
2. The Keel website, https://sci2s.ugr.es/keel/datasets.php(accessed on 24 November 2024)[2].
3. The Statlib URL ftp://lib.stat.cmu.edu/datasets/index.html(accessed on 24 November 2024).

3.1. Datasets

The following classification datasets were incorporated in the conducted experiments:

1. **Alcohol**. This dataset is related to experiments on alcohol consumption [3].
2. **Australian**, related to economic risk in bank transactions [4].
3. **Bands**, related to printing problems [5], with two distinct classes.
4. **Cleveland**, a medical dataset studied in many research papers [6,7].
5. **Circular** dataset, which is an artificial dataset.
6. **Dermatology**, a medical dataset related to dermatology problems [8] with 6 classes.
7. **Ecoli**, that is related to issues about proteins [9].
8. **Haberman**, a medical dataset used in the detection of breast cancer.
9. **Hayes-roth** dataset [10], a dataset with 3 classes.
10. **Heart**, a medical dataset about heart diseases [11] with two classes.
11. **HeartAttack**, a medical dataset related to the presence of heart diseases, with two classes.
12. **Hepatitis**, a dataset used for to detect the presence of hepatitis.
13. **Housevotes**, that uses data from the Congressional voting in USA [12].
14. **Ionosphere**, that used data from various measurements in the ionosphere [13,14].
15. **Liverdisorder**, which is a medical dataset [15,16] with two classes.
16. **Lymography** [17], which is a dataset with four distinct classes.
17. **Magic**, this dataset contains data from simulation regarding gamma particles [18].
18. **Mammographic**, a medical dataset used for the detection of breast cancer [19].
19. **Page Blocks** dataset [23], which is incorporated for the detection of page layout in documents.
20. **Parkinsons**, a dataset used to detect the presence of Parkinson's disease [20,21].
21. **Pima**, a medical dataset that used to detect the presence of diabetes[22].
22. **Phoneme**, where the purpose of this dataset is to distinguish between nasal and oral sounds.
23. **Popfailures**, a dataset related to climate measurements [24].
24. **Regions2**, used detect issues in the liver from various liver biopsy images [25].
25. **Ring**, a dataset related to some multivariate normal distributions.
26. **Saheart**, used to detect the presence of some heart diseases.[26].
27. **Segment** dataset [27], which is used for image processing .
28. **Statheart**, a medical dataset about heart diseases.
29. **Spambase**, a dataset used for the detection of spam emails.
30. **Spiral**, which is an artificial dataset.
31. **Student**, a dataset related to some experiments in schools [28].
32. **Tae**, this dataset is related to evaluations of teaching performance.
33. **Transfusion**, which is a medical dataset [29].

34. **Wdbc**, a medical dataset related to the detection of breast cancer [30,31].
35. **Wine**, a dataset related to the quality of wines [32,33].
36. **EEG** dataset, which is a medical dataset related to EEG measurements[34,35]. From the original dataset the following cases were obtained: Z_F_S, ZO_NF_S, Z_O_N_F_S and ZONF_S.
37. **Zoo**, that used to predict the class of some animals [36] .

Also the following regression datasets were used in the conducted experiments:

1. **Abalone**, a dataset regarding the prediction of the age of abalones [37].
2. **Airfoil**, a dataset derived from NASA [38] with 5 features.
3. **Auto**, a dataset used for the prediction of fuel consumption in cars.
4. **BK**, related to basketball games. The dataset has 4 features.
5. **BL**, a dataset used in some electricity experiments and it contains 7 features.
6. **Baseball**, a dataset with 16 features used for the estimation of the income of baseball players.
7. **Concrete**, a dataset with 8 features and it was used in civil engineering [39].
8. **DEE**, a dataset with 6 features, used in the prediction of electricity prices.
9. **EU**, a dataset originated in the STATLIB website.
10. **FA**, that contains related to body fat.
11. **Friedman**, a synthetic dataset used in various benchmarks [40].
12. **FY,** this dataset used to measure the longevity of fruit flies.
13. **HO**, a dataset founded in the STATLIB repository with 13 features.
14. **Housing**, used to predict the price of houses [41].
15. **Laser**, which is a dataset with 4 features and it has been used in various laser experiments.
16. **LW**, a dataset with 9 features used to measure the weight of babes.
17. **Mortgage**, an economic dataset with 15 features.
18. **Plastic**, a dataset related to the detection of pressure on plastics.
19. **PL**, a dataset with 2 features founded in the STATLIB repository.
20. **RealEstate,** a dataset found in the STATLIB website with 5 features.
21. **Quake**, a dataset used in earthquake measurements with 3 features.
22. **SN**, a dataset that provides experimental measurements related to trellising and pruning, with 11 features.
23. **Stock**, a dataset with 9 features for the prediction of the prices of various stocks.
24. **Treasury**, an economic dataset with 15 features.
25. **TZ**, derived from the STATLIB website and it has 60 features.
26. **VE** dataset, derived from the STALIB repository.

3.2. Experimental results 101

Table 1. Experimental results for the used classification datasets. The numbers in cells represent average classification error as measured on the corresponding test set.

DATASET	ADAM	BFGS	GENETIC	RBF	PROPOSED
APPENDICITIS	16.50%	18.00%	24.40%	12.23%	15.00%
ALCOHOL	57.78%	41.50%	39.57%	49.32%	18.33%
AUSTRALIAN	35.65%	38.13%	32.21%	34.89%	21.49%
BALANCE	12.27%	8.64%	8.97%	33.53%	7.79%
CLEVELAND	67.55%	77.55%	51.60%	67.10%	42.38%
CIRCULAR	19.95%	6.08%	5.99%	5.98%	6.50%
DERMATOLOGY	26.14%	52.92%	30.58%	62.34%	4.97%
ECOLI	64.43%	69.52%	54.67%	59.48%	40.30%
GLASS	61.38%	54.67%	52.86%	50.46%	54.38%
HABERMAN	29.00%	29.34%	28.66%	25.10%	26.53%
HAYES-ROTH	59.70%	37.33%	56.18%	64.36%	34.31%
HEART	38.53%	39.44%	28.34%	31.20%	13.11%
HEARTATTACK	45.55%	46.67%	29.03%	29.00%	21.90%
HOUSEVOTES	7.48%	7.13%	6.62%	6.13%	6.09%
IONOSPHERE	16.64%	15.29%	15.14%	16.22%	10.37%
LIVERDISORDER	41.53%	42.59%	31.11%	30.84%	29.94%
LYMOGRAPHY	39.79%	35.43%	28.42%	25.50%	17.93%
MAMMOGRAPHIC	46.25%	17.24%	19.88%	21.38%	16.63%
PARKINSONS	24.06%	27.58%	18.05%	17.41%	12.79%
PHONEME	29.43%	15.58%	15.55%	23.32%	18.10%
PIMA	34.85%	35.59%	32.19%	25.78%	25.03%
POPFAILURES	5.18%	5.24%	5.94%	7.04%	4.45%
REGIONS2	29.85%	36.28%	29.39%	38.29%	25.19%
SAHEART	34.04%	37.48%	34.86%	32.19%	29.26%
SEGMENT	49.75%	68.97%	57.72%	59.68%	27.80%
SONAR	30.33%	25.85%	22.40%	27.90%	20.50%
SPIRAL	47.67%	47.99%	48.66%	44.87%	41.60%
STATHEART	44.04%	39.65%	27.25%	31.36%	19.74%
STUDENT	5.13%	7.14%	5.61%	5.49%	4.00%
TRANSFUSION	25.68%	25.84%	24.87%	26.41%	23.35%
WDBC	35.35%	29.91%	8.56%	7.27%	6.73%
WINE	29.40%	59.71%	19.20%	31.41%	6.29%
Z_F_S	47.81%	39.37%	10.73%	13.16%	8.38%
ZO_NF_S	47.43%	43.04%	21.54%	9.02%	4.32%
ZONF_S	11.99%	15.62%	4.36%	4.03%	1.76%
ZOO	14.13%	10.70%	9.50%	21.93%	7.00%
AVERAGE	**34.23%**	**33.58%**	**26.13%**	**29.21%**	**18.73%**

Table 2. Experimental results for the used regression datasets. Numbers in cells represent average regression error as calculated on the corresponding test set.

DATASET	ADAM	BFGS	GENETIC	RBF	PROPOSED
ABALONE	4.30	5.69	7.17	7.37	4.32
AIRFOIL	0.005	0.003	0.003	0.27	0.002
AUTO	70.84	60.97	12.18	17.87	12.78
BK	0.0252	0.28	0.027	0.02	0.02
BL	0.622	2.55	5.74	0.013	0.006
BASEBALL	77.90	119.63	103.60	93.02	60.74
CONCRETE	0.078	0.066	0.0099	0.011	0.006
DEE	0.63	2.36	1.013	0.17	0.19
FRIEDMAN	22.90	1.263	1.249	7.23	2.21
FY	0.038	0.19	0.65	0.041	0.067
HO	0.035	0.62	2.78	0.03	0.015
HOUSING	80.99	97.38	43.26	57.68	20.74
LASER	0.03	0.015	0.59	0.03	0.004
LW	0.028	2.98	1.90	0.03	0.011
MORTGAGE	9.24	8.23	2.41	1.45	0.32
PL	0.117	0.29	0.29	2.118	0.022
PLASTIC	11.71	20.32	2.791	8.62	2.16
QUAKE	0.07	0.42	0.04	0.07	0.036
SN	0.026	0.40	2.95	0.027	0.023
STOCK	180.89	302.43	3.88	12.23	5.57
TREASURY	11.16	9.91	2.93	2.02	0.68
AVERAGE	**22.46**	**30.29**	**9.31**	**10.02**	**5.23**

4. Conclusions

Author Contributions: V.C. and I.G.T. conducted the experiments, employing several datasets and provided the comparative experiments. D.T. and V.C. performed the statistical analysis and prepared the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This research has been financed by the European Union : Next Generation EU through the Program Greece 2.0 National Recovery and Resilience Plan , under the call RESEARCH – CREATE – INNOVATE, project name “iCREW: Intelligent small craft simulator for advanced crew training using Virtual Reality techniques" (project code:TAEDK-06195).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. M. Kelly, R. Longjohn, K. Nottingham, The UCI Machine Learning Repository, https://archive.ics.uci.edu.
2. J. Alcalá-Fdez, A. Fernandez, J. Luengo, J. Derrac, S. García, L. Sánchez, F. Herrera. KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework. Journal of Multiple-Valued Logic and Soft Computing 17, pp. 255-287, 2011.
3. Tzimourta, K.D.; Tsoulos, I.; Bilero, I.T.; Tzallas, A.T.; Tsipouras, M.G.; Giannakeas, N. Direct Assessment of Alcohol Consumption in Mental State Using Brain Computer Interfaces and Grammatical Evolution. Inventions 2018, 3, 51.
4. J.R. Quinlan, Simplifying Decision Trees. International Journal of Man-Machine Studies **27**, pp. 221-234, 1987.
5. B. Evans, D. Fisher, Overcoming process delays with decision tree induction. IEEE Expert **9**, pp. 60-66, 1994.
6. Z.H. Zhou,Y. Jiang, NeC4.5: neural ensemble based C4.5," in IEEE Transactions on Knowledge and Data Engineering **16**, pp. 770-773, 2004.

7. R. Setiono , W.K. Leow, FERNN: An Algorithm for Fast Extraction of Rules from Neural Networks, Applied Intelligence **12**, pp. 15-25, 2000.
8. G. Demiroz, H.A. Govenir, N. Ilter, Learning Differential Diagnosis of Eryhemato-Squamous Diseases using Voting Feature Intervals, Artificial Intelligence in Medicine. **13**, pp. 147–165, 1998.
9. P. Horton, K.Nakai, A Probabilistic Classification System for Predicting the Cellular Localization Sites of Proteins, In: Proceedings of International Conference on Intelligent Systems for Molecular Biology **4**, pp. 109-15, 1996.
10. B. Hayes-Roth, B., F. Hayes-Roth. Concept learning and the recognition and classification of exemplars. Journal of Verbal Learning and Verbal Behavior **16**, pp. 321-338, 1977.
11. I. Kononenko, E. Šimec, M. Robnik-Šikonja, Overcoming the Myopia of Inductive Learning Algorithms with RELIEFF, Applied Intelligence **7**, pp. 39–55, 1997
12. R.M. French, N. Chater, Using noise to compute error surfaces in connectionist networks: a novel means of reducing catastrophic forgetting, Neural Comput. **14**, pp. 1755-1769, 2002.
13. J.G. Dy , C.E. Brodley, Feature Selection for Unsupervised Learning, The Journal of Machine Learning Research **5**, pp 845–889, 2004.
14. S. J. Perantonis, V. Virvilis, Input Feature Extraction for Multilayered Perceptrons Using Supervised Principal Component Analysis, Neural Processing Letters **10**, pp 243–252, 1999.
15. J. Garcke, M. Griebel, Classification with sparse grids using simplicial basis functions, Intell. Data Anal. **6**, pp. 483-502, 2002.
16. J. Mcdermott, R.S. Forsyth, Diagnosing a disorder in a classification benchmark, Pattern Recognition Letters **73**, pp. 41-43, 2016.
17. G. Cestnik, I. Konenenko, I. Bratko, Assistant-86: A Knowledge-Elicitation Tool for Sophisticated Users. In: Bratko, I. and Lavrac, N., Eds., Progress in Machine Learning, Sigma Press, Wilmslow, pp. 31-45, 1987.
18. Heck, D., Knapp, J., Capdevielle, J. N., Schatz, G., & Thouw, T. (1998). CORSIKA: A Monte Carlo code to simulate extensive air showers.
19. M. Elter, R. Schulz-Wendtland, T. Wittenberg, The prediction of breast cancer biopsy outcomes using two CAD approaches that both emphasize an intelligible decision process, Med Phys. **34**, pp. 4164-72, 2007.
20. M.A. Little, P.E. McSharry, S.J Roberts et al, Exploiting Nonlinear Recurrence and Fractal Scaling Properties for Voice Disorder Detection. BioMed Eng OnLine **6**, 23, 2007.
21. M.A. Little, P.E. McSharry, E.J. Hunter, J. Spielman, L.O. Ramig, Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE Trans Biomed Eng. **56**, pp. 1015-1022, 2009.
22. J.W. Smith, J.E. Everhart, W.C. Dickson, W.C. Knowler, R.S. Johannes, Using the ADAP learning algorithm to forecast the onset of diabetes mellitus, In: Proceedings of the Symposium on Computer Applications and Medical Care IEEE Computer Society Press, pp.261-265, 1988.
23. F. Esposito F., D. Malerba, G. Semeraro, Multistrategy Learning for Document Recognition, Applied Artificial Intelligence **8**, pp. 33-84, 1994.
24. D.D. Lucas, R. Klein, J. Tannahill, D. Ivanova, S. Brandon, D. Domyancic, Y. Zhang, Failure analysis of parameter-induced simulation crashes in climate models, Geoscientific Model Development **6**, pp. 1157-1171, 2013.
25. N. Giannakeas, M.G. Tsipouras, A.T. Tzallas, K. Kyriakidi, Z.E. Tsianou, P. Manousou, A. Hall, E.C. Karvounis, V. Tsianos, E. Tsianos, A clustering based method for collagen proportional area extraction in liver biopsy images (2015) Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2015-November, art. no. 7319047, pp. 3097-3100.
26. T. Hastie, R. Tibshirani, Non-parametric logistic and proportional odds regression, JRSS-C (Applied Statistics) **36**, pp. 260–276, 1987.
27. M. Dash, H. Liu, P. Scheuermann, K. L. Tan, Fast hierarchical clustering and its validation, Data & Knowledge Engineering **44**, pp 109–138, 2003.
28. P. Cortez, A. M. Gonçalves Silva, Using data mining to predict secondary school student performance, In Proceedings of 5th FUture BUsiness TEChnology Conference (FUBUTEC 2008) (pp. 5–12). EUROSIS-ETI, 2008.
29. I-Cheng Yeh, King-Jang Yang, Tao-Ming Ting, Knowledge discovery on RFM model using Bernoulli sequence, Expert Systems with Applications **36**, pp. 5866-5871, 2009.
30. Jeyasingh, S., & Veluchamy, M. (2017). Modified bat algorithm for feature selection with the wisconsin diagnosis breast cancer (WDBC) dataset. Asian Pacific journal of cancer prevention: APJCP, 18(5), 1257.
31. Alshayeji, M. H., Ellethy, H., & Gupta, R. (2022). Computer-aided detection of breast cancer on the Wisconsin dataset: An artificial neural networks approach. Biomedical signal processing and control, 71, 103141.
32. M. Raymer, T.E. Doom, L.A. Kuhn, W.F. Punch, Knowledge discovery in medical and biological datasets using a hybrid Bayes classifier/evolutionary algorithm. IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics : a publication of the IEEE Systems, Man, and Cybernetics Society, **33** , pp. 802-813, 2003.
33. P. Zhong, M. Fukushima, Regularized nonsmooth Newton method for multi-class support vector machines, Optimization Methods and Software **22**, pp. 225-236, 2007.
34. R. G. Andrzejak, K. Lehnertz, F.Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state," Physical Review E, vol. 64, no. 6, Article ID 061907, 8 pages, 2001.

35. A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Automatic Seizure Detection Based on Time-Frequency Analysis and Artificial Neural Networks," Computational Intelligence and Neuroscience, vol. 2007, Article ID 80510, 13 pages, 2007. doi:10.1155/2007/80510
36. M. Koivisto, K. Sood, Exact Bayesian Structure Discovery in Bayesian Networks, The Journal of Machine Learning Research **5**, pp. 549–573, 2004.
37. Nash, W.J.; Sellers, T.L.; Talbot, S.R.; Cawthor, A.J.; Ford, W.B. The Population Biology of Abalone (_Haliotis_ species) in Tasmania. I. Blacklip Abalone (_H. rubra_) from the North Coast and Islands of Bass Strait, Sea Fisheries Division; Technical Report No. 48; Department of Primary Industry and Fisheries, Tasmania: Hobart, Australia, 1994; ISSN 1034-3288
38. Brooks, T.F.; Pope, D.S.; Marcolini, A.M. Airfoil Self-Noise and Prediction. Technical Report, NASA RP-1218. July 1989. Available online: https://ntrs.nasa.gov/citations/19890016302 (accessed on 14 November 2024).
39. I.Cheng Yeh, Modeling of strength of high performance concrete using artificial neural networks, Cement and Concrete Research. **28**, pp. 1797-1808, 1998.
40. Friedman, J. (1991): Multivariate Adaptative Regression Splines. Annals of Statistics, 19:1, 1--141.
41. D. Harrison and D.L. Rubinfeld, Hedonic prices and the demand for clean ai, J. Environ. Economics & Management **5**, pp. 81-102, 1978.
42. K. O. Stanley, R. Miikkulainen, Evolving Neural Networks through Augmenting Topologies, Evolutionary Computation **10**, pp. 99-127, 2002.
43. J. Park and I. W. Sandberg, Universal Approximation Using Radial-Basis-Function Networks, Neural Computation **3**, pp. 246-257, 1991.
44. G.A. Montazer, D. Giveki, M. Karami, H. Rastegar, Radial basis function neural networks: A review. Comput. Rev. J **1**, pp. 52-74, 2018.