Article

Neural DE: an evolutionary method based on Differential Evolution suitable for neural network training

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

[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
* Correspondence: itsoulos@uoi.gr;

Abstract: Artificial neural networks have proven to be an important machine learning model that has been widely used in recent decades in a number of difficult problems of classification or data fitting from real - world areas. Due to their importance, a number of techniques have been developed that efficiently identify the parameter vector for these models. These techniques usually come from the field of optimization and, by minimizing the training error of artificial neural networks, estimate the vector of their parameters. However, many times these techniques either get trapped in local minima of the training error or lead to overfitting the artificial neural network, resulting in poor performance when applied to data that was not present during the training process. This paper presents an innovative training technique for artificial neural networks based on the Differential Evolution optimization method. This new technique creates an initial population of artificial neural networks that evolve and periodically applies a local optimization technique in order to accelerate the training of these networks. The application of the local minimization technique is done in such a way as to avoid the phenomenon of overfitting. This new method was successfully applied to a series of classification and data fitting problems and a comparative study was made with other training techniques from the relevant literature.

Keywords: Neural networks; Evolutionary methods; Differential Evolution; Machine learning

Citation: Tsoulos, I.G.; Charilogis V. Neural DE: an evolutionary method based on Differential Evolution suitable for neural network training. *Journal Not Specified* **2022**, *1*, 0. https://doi.org/

Received:
Accepted:
Published:

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

2. The proposed method

3. Experiments

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. Experimental 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. **FA**, that contains related to body fat.
10. **Friedman**, a synthetic dataset used in various benchmarks [40].
11. **FY,** this dataset used to measure the longevity of fruit flies.
12. **HO**, a dataset founded in the STATLIB repository with 13 features.
13. **Housing**, used to predict the price of houses [41].
14. **Laser**, which is a dataset with 4 features and it has been used in various laser experiments.
15. **LW**, a dataset with 9 features used to measure the weight of babes.
16. **Mortgage**, an economic dataset with 15 features.
17. **Plastic**, a dataset related to the detection of pressure on plastics.
18. **PL**, a dataset with 2 features founded in the STATLIB repository.
19. **RealEstate,** a dataset found in the STATLIB website with 5 features.
20. **Quake**, a dataset used in earthquake measurements with 3 features.
21. **SN**, a dataset that provides experimental measurements related to trellising and pruning, with 11 features.
22. **Stock**, a dataset with 9 features for the prediction of the prices of various stocks.
23. **Treasury**, an economic dataset with 15 features.

3.2. Experimental results

Table 1. Experimental results for the classification datasets using the series of optimization methods. The numbers in cells represent average classification error measured on the corresponding test set.

DATASET	ADAM	BFGS	GENETIC	NEAT	RBF	NEURALDE
ALCOHOL	57.78%	41.50%	39.57%	66.80%	49.32%	19.15%
AUSTRALIAN	35.65%	38.13%	32.21%	31.98%	34.89%	15.31%
BALANCE	12.27%	8.64%	8.97%	23.14%	33.53%	6.92%
BANDS	36.92%	36.67%	34.92%	34.30%	37.17%	35.00%
CLEVELAND	67.55%	77.55%	51.60%	53.44%	67.10%	45.07%
CIRCULAR	19.95%	6.08%	5.99%	35.18%	5.98%	4.23%
DERMATOLOGY	26.14%	52.92%	30.58%	32.43%	62.34%	10.38%
ECOLI	64.43%	69.52%	54.67%	43.44%	59.48%	45.22%
HABERMAN	29.00%	29.34%	28.66%	24.04%	25.10%	27.55%
HAYES-ROTH	59.70%	37.33%	56.18%	50.15%	64.36%	35.59%
HEART	38.53%	39.44%	28.34%	39.27%	31.20%	17.60%
HEARTATTACK	45.55%	46.67%	29.03%	32.34%	29.00%	19.70%
HEPATITIS	68.13%	72.47%	62.12%	67.04%	64.63%	57.46%
HOUSEVOTES	7.48%	7.13%	6.62%	10.89%	6.13%	7.48%
IONOSPHERE	16.64%	15.29%	15.14%	19.67%	16.22%	16.17%
LIVERDISORDER	41.53%	42.59%	31.11%	30.67%	30.84%	31.72%
LYMOGRAPHY	39.79%	35.43%	28.42%	33.70%	25.50%	28.86%
MAGIC	40.55%	17.30%	21.75%	24.85%	21.28%	11.73%
MAMMOGRAPHIC	46.25%	17.24%	19.88%	22.85%	21.38%	17.52%
PARKINSONS	24.06%	27.58%	18.05%	18.56%	17.41%	14.32%
PAGE BLOCKS	34.27%	8.49%	6.84%	10.22%	10.09%	6.04%
PHONEME	29.43%	15.58%	15.55%	22.34%	23.32%	15.50%
PIMA	34.85%	35.59%	32.19%	34.51%	25.78%	24.85%
POPFAILURES	5.18%	5.24%	5.94%	7.05%	7.04%	6.09%
REGIONS2	29.85%	36.28%	29.39%	33.23%	38.29%	28.77%
RING	28.80%	29.44%	28.80%	30.85%	21.67%	22.90%
SAHEART	34.04%	37.48%	34.86%	34.51%	32.19%	29.63%
SEGMENT	49.75%	68.97%	57.72%	66.72%	59.68%	15.60%
SONAR	30.33%	25.85%	22.40%	34.10%	27.90%	19.80%
SPAMBASE	48.05%	18.16%	6.37%	35.77%	29.35%	4.95%
SPIRAL	47.67%	47.99%	48.66%	48.66%	44.87%	42.06%
STATHEART	44.04%	39.65%	27.25%	44.36%	31.36%	18.53%
STUDENT	5.13%	7.14%	5.61%	10.20%	5.49%	4.86%
TAE	60.20%	51.58%	49.84%	60.67%	60.02%	45.62%
TRANSFUSION	25.68%	25.84%	24.87%	24.87%	26.41%	23.59%
WDBC	35.35%	29.91%	8.56%	12.88%	7.27%	4.29%
WINE	29.40%	59.71%	19.20%	25.43%	31.41%	10.39%
Z_F_S	47.81%	39.37%	10.73%	38.41%	13.16%	6.81%
ZO_NF_S	47.43%	43.04%	21.54%	43.75%	9.02%	4.73%
ZONF_S	11.99%	15.62%	4.36%	5.44%	4.03%	2.41%
ZOO	14.13%	10.70%	9.50%	20.27%	21.93%	6.63%
AVERAGE	**35.88%**	**33.43%**	**26.19%**	**32.66%**	**30.08%**	**19.78%**

Table 2. Experimental results for the provided regression datasets using the series of methods. Numbers in cells represents average regression error calculated on the corresponding test set.

DATASET	ADAM	BFGS	GENETIC	NEAT	RBF	NEURALDE
ABALONE	4.30	5.69	7.17	9.88	7.37	5.04
AIRFOIL	0.005	0.003	0.003	0.067	0.27	0.0014
AUTO	70.84	60.97	12.18	56.06	17.87	16.21
BK	0.0252	0.28	0.027	0.15	0.02	0.019
BL	0.622	2.55	5.74	0.05	0.013	0.016
BASEBALL	77.90	119.63	103.60	100.39	93.02	60.56
CONCRETE	0.078	0.066	0.0099	0.081	0.011	0.003
DEE	0.63	2.36	1.013	1.51	0.17	0.35
FA	0.048	0.426	0.025	0.19	0.015	0.083
FRIEDMAN	22.90	1.263	1.249	19.35	7.23	1.22
HO	0.035	0.62	2.78	0.17	0.03	0.017
HOUSING	80.99	97.38	43.26	56.49	57.68	24.82
LASER	0.03	0.015	0.59	0.084	0.03	0.0026
LW	0.028	2.98	1.90	0.17	0.03	0.021
MORTGAGE	9.24	8.23	2.41	14.11	1.45	0.54
PLASTIC	11.71	20.32	2.791	20.77	8.62	3.27
PY	0.321	0.578	0.56	0.075	0.012	0.11
QUAKE	0.07	0.42	0.04	0.298	0.07	0.042
SN	0.026	0.40	2.95	0.174	0.027	0.027
STOCK	180.89	302.43	3.88	215.82	12.23	3.40
TREASURY	11.16	9.91	2.93	15.52	2.02	1.021
AVERAGE	**22.47**	**30.31**	**9.29**	**24.35**	**9.91**	**5.56**

3.2.1. Experiments with the differential weight

Table 3. Experimental results for the classification datasets using a list of proposed differential weight mechanisms found in the literature.

DATASET	MIGRANT	ADAPTIVE	RANDOM
ALCOHOL	19.62%	18.44%	19.15%
AUSTRALIAN	14.77%	16.14%	15.31%
BALANCE	6.79%	6.94%	6.92%
BANDS	33.28%	35.42%	35.00%
CLEVELAND	45.74%	44.60%	45.07%
CIRCULAR	4.28%	4.34%	4.23%
DERMATOLOGY	11.41%	9.83%	10.38%
ECOLI	42.61%	47.51%	45.22%
HABERMAN	27.76%	28.35%	27.55%
HAYES-ROTH	35.92%	36.31%	35.59%
HEART	19.01%	17.29%	17.60%
HEARTATTACK	21.31%	20.28%	19.70%
HEPATITIS	56.37%	56.46%	57.46%
HOUSEVOTES	7.47%	7.54%	7.48%
IONOSPHERE	15.88%	16.40%	16.17%
LIVERDISORDER	32.56%	32.54%	31.72%
LYMOGRAPHY	26.52%	28.95%	28.86%
MAGIC	11.83%	11.28%	11.73%
MAMMOGRAPHIC	17.69%	17.44%	17.52%
PARKINSONS	13.28%	14.21%	14.32%
PAGE BLOCKS	5.48%	5.93%	6.04%
PHONEME	15.39%	15.17%	15.50%
PIMA	25.04%	23.28%	24.85%
POPFAILURES	5.55%	6.55%	6.09%
REGIONS2	26.26%	28.98%	28.77%
RING	21.04%	25.54%	22.90%
SAHEART	29.94%	29.80%	29.63%
SEGMENT	9.71%	15.99%	15.60%
SONAR	18.83%	18.93%	19.80%
SPAMBASE	5.87%	4.78%	4.95%
SPIRAL	41.36%	41.54%	42.06%
STATHEART	20.79%	19.32%	18.53%
STUDENT	4.36%	5.28%	4.86%
TAE	45.09%	44.78%	45.62%
TRANSFUSION	23.01%	24.51%	23.59%
WDBC	3.80%	4.55%	4.29%
WINE	8.22%	12.84%	10.39%
Z_F_S	7.16%	7.15%	6.81%
ZO_NF_S	4.19%	4.70%	4.73%
ZONF_S	2.47%	2.44%	2.41%
ZOO	5.23%	6.87%	6.63%
AVERAGE	**19.34%**	**19.98%**	**19.78%**

Table 4. Experimental results for the regression datasets using a variety of differential weigh methods.

DATASET	MIGRANT	ADAPTIVE	RANDOM
ABALONE	4.57	5.97	5.04
AIRFOIL	0.002	0.0011	0.0014
AUTO	45.37	18.89	16.21
BK	0.28	0.02	0.019
BL	0.007	0.019	0.016
BASEBALL	79.24	61.89	60.56
CONCRETE	0.0028	0.0029	0.003
DEE	0.27	0.35	0.35
FA	0.051	0.10	0.083
FRIEDMAN	1.66	1.32	1.22
HO	0.009	0.014	0.017
HOUSING	14.58	32.62	24.82
LASER	0.0026	0.0027	0.0026
LW	0.0189	0.044	0.021
MORTGAGE	0.43	0.93	0.54
PLASTIC	5.36	3.57	3.27
PY	0.12	0.14	0.11
QUAKE	0.038	0.041	0.042
SN	0.023	0.031	0.027
STOCK	15.94	3.58	3.40
TREASURY	0.70	1.24	1.021
AVERAGE	**8.03**	**6.23**	**5.56**

3.2.2. Experiments with the factor a 101

Table 5. Experimental results for the classification datasets using different values for the critical parameter a

DATASET	$a = 1.25$	$a = 1.5$	$a = 2$	$a = 4$	$a = 8$
ALCOHOL	21.49%	20.20%	19.15%	22.34%	19.86%
AUSTRALIAN	16.43%	16.89%	15.31%	15.90%	16.16%
BALANCE	7.49%	7.56%	6.92%	7.01%	7.16%
BANDS	34.88%	35.09%	35.00%	35.61%	35.01%
CLEVELAND	44.86%	46.08%	45.07%	44.67%	45.07%
CIRCULAR	4.17%	4.26%	4.23%	4.49%	4.61%
DERMATOLOGY	10.77%	10.65%	10.38%	9.87%	8.94%
ECOLI	44.13%	45.87%	45.22%	46.65%	47.46%
HABERMAN	27.24%	27.02%	27.55%	27.91%	28.40%
HAYES-ROTH	36.95%	36.46%	35.59%	35.82%	33.21%
HEART	17.43%	17.17%	17.60%	17.19%	17.19%
HEARTATTACK	19.50%	19.27%	19.70%	19.69%	19.69%
HEPATITIS	59.00%	57.38%	57.46%	60.21%	60.96%
HOUSEVOTES	7.07%	7.90%	7.48%	7.06%	7.07%
IONOSPHERE	14.24%	15.48%	16.17%	16.17%	15.43%
LIVERDISORDER	31.41%	31.59%	31.72%	32.23%	32.33%
LYMOGRAPHY	27.00%	27.07%	28.86%	27.57%	27.09%
MAGIC	12.18%	11.94%	11.73%	11.78%	12.25%
MAMMOGRAPHIC	17.42%	17.46%	17.52%	17.56%	17.54%
PARKINSONS	13.67%	13.40%	14.32%	14.19%	13.83%
PAGE BLOCKS	6.32%	6.28%	6.04%	5.83%	5.92%
PHONEME	16.07%	15.69%	15.50%	15.00%	15.00%
PIMA	24.59%	24.61%	24.85%	24.57%	24.79%
POPFAILURES	4.90%	5.28%	6.09%	7.50%	6.94%
REGIONS2	29.35%	28.94%	28.77%	28.79%	28.69%
RING	27.65%	26.97%	22.90%	23.92%	23.20%
SAHEART	29.33%	29.37%	29.63%	29.97%	30.47%
SEGMENT	18.18%	16.86%	15.60%	14.53%	14.27%
SONAR	19.27%	19.62%	19.80%	18.87%	20.16%
SPAMBASE	4.67%	5.07%	4.95%	5.00%	5.87%
SPIRAL	41.92%	41.76%	42.06%	41.49%	42.13%
STATHEART	19.22%	18.91%	18.53%	19.30%	20.22%
STUDENT	4.76%	4.55%	4.86%	5.84%	6.43%
TAE	47.54%	46.22%	45.62%	46.07%	44.38%
TRANSFUSION	23.74%	23.87%	23.59%	23.95%	24.15%
WDBC	4.48%	4.50%	4.29%	4.12%	4.14%
WINE	9.24%	9.67%	10.39%	10.39%	11.88%
Z_F_S	7.26%	6.67%	6.81%	6.64%	6.70%
ZO_NF_S	4.82%	4.91%	4.73%	4.42%	4.33%
ZONF_S	2.42%	2.55%	2.41%	2.68%	2.60%
ZOO	6.23%	6.93%	6.63%	7.07%	6.40%
AVERAGE	**19.98%**	**19.95%**	**19.78%**	**20.00%**	**19.95%**

Table 6. Experimental results for the regression datasets using different values for the critical parameter a.

DATASET	$a = 1.25$	$a = 1.5$	$a = 2$	$a = 4$	$a = 8$
ABALONE	4.36	4.66	5.04	7.63	8.84
AIRFOIL	0.002	0.0018	0.0014	0.0009	0.0007
AUTO	14.97	14.72	16.21	16.04	16.58
BK	0.02	0.02	0.019	0.11	0.21
BL	0.011	0.012	0.016	0.037	0.11
BASEBALL	57.67	58.42	60.56	67.70	72.37
CONCRETE	0.003	0.003	0.003	0.004	0.005
DEE	0.35	0.34	0.35	0.34	0.34
FA	0.067	0.056	0.083	0.24	0.42
FRIEDMAN	1.30	1.24	1.22	1.41	19.74
HO	0.011	0.012	0.017	0.043	0.17
HOUSING	28.09	25.93	24.82	22.86	26.45
LASER	0.0026	0.0025	0.0026	0.0027	0.0035
LW	0.019	0.019	0.021	0.073	0.77
MORTGAGE	0.50	0.45	0.54	0.75	0.62
PLASTIC	3.37	3.10	3.27	3.02	2.80
PY	0.12	0.13	0.11	0.09	0.09
QUAKE	0.04	0.041	0.042	0.048	0.096
SN	0.024	0.024	0.027	0.087	0.17
STOCK	2.82	3.13	3.40	3.21	5.42
TREASURY	0.87	0.81	1.021	1.06	1.06
AVERAGE	**5.46**	**5.39**	**5.56**	**5.94**	**7.44**

3.2.3. Experiments with the local search rate 102

Table 7. Experimental results for the classification datasets using different values of local search rate p_l.

DATASET	$p_l = 0.0025$	$p_l = 0.005$	$p_l = 0.01$	$p_l = 0.02$
ALCOHOL	19.15%	19.15%	21.81%	16.49%
AUSTRALIAN	17.93%	15.31%	14.17%	14.14%
BALANCE	7.44%	6.92%	6.63%	6.54%
BANDS	35.66%	35.00%	35.01%	34.44%
CLEVELAND	46.45%	45.07%	45.39%	45.32%
CIRCULAR	4.21%	4.23%	4.16%	4.31%
DERMATOLOGY	11.25%	10.38%	9.96%	9.07%
ECOLI	46.64%	45.22%	44.34%	44.66%
HABERMAN	27.99%	27.55%	27.66%	27.43%
HAYES-ROTH	36.33%	35.59%	34.03%	31.44%
HEART	17.05%	17.60%	17.35%	17.51%
HEARTATTACK	19.62%	19.70%	19.43%	20.37%
HEPATITIS	60.13%	57.46%	58.79%	60.34%
HOUSEVOTES	7.81%	7.48%	8.02%	7.47%
IONOSPHERE	15.52%	16.17%	16.33%	16.72%
LIVERDISORDER	31.92%	31.72%	32.06%	32.40%
LYMOGRAPHY	27.48%	28.86%	29.55%	29.98%
MAGIC	12.00%	11.73%	11.55%	11.57%
MAMMOGRAPHIC	17.23%	17.52%	17.34%	17.57%
PARKINSONS	14.70%	14.32%	13.39%	14.04%
PAGE BLOCKS	6.02%	6.04%	5.66%	5.76%
PHONEME	15.67%	15.50%	15.29%	15.04%
PIMA	25.52%	24.85%	24.63%	24.87%
POPFAILURES	5.46%	6.09%	6.30%	6.75%
REGIONS2	30.41%	28.77%	28.77%	28.02%
RING	27.01%	22.90%	24.05%	23.85%
SAHEART	29.91%	29.63%	29.54%	29.89%
SEGMENT	19.36%	15.60%	13.88%	11.70%
SONAR	19.62%	19.80%	20.38%	18.53%
SPAMBASE	6.30%	4.95%	5.43%	4.78%
SPIRAL	42.22%	42.06%	41.16%	40.35%
STATHEART	19.90%	18.53%	19.72%	20.47%
STUDENT	4.97%	4.86%	5.13%	5.01%
TAE	47.05%	45.62%	43.93%	45.05%
TRANSFUSION	24.41%	23.59%	23.11%	22.64%
WDBC	4.88%	4.29%	4.17%	4.07%
WINE	12.39%	10.39%	9.02%	8.41%
Z_F_S	7.11%	6.81%	6.26%	6.62%
ZO_NF_S	5.69%	4.73%	4.23%	3.93%
ZONF_S	2.45%	2.41%	2.45%	2.38%
ZOO	6.93%	6.63%	6.00%	6.50%
AVERAGE	**20.48%**	**19.78%**	**19.66%**	**19.43%**

Table 8. Experimental results for the regression datasets using different values of the local search rate parameter p_l.

DATASET	$p_l = 0.0025$	$p_l = 0.005$	$p_l = 0.01$	$p_l = 0.02$
ABALONE	4.48	5.04	5.27	4.93
AIRFOIL	0.0019	0.0014	0.0009	0.0006
AUTO	17.63	16.21	12.50	10.11
BK	0.027	0.019	0.03	0.029
BL	0.031	0.016	0.05	0.017
BASEBALL	63.05	60.56	62.57	71.93
CONCRETE	0.003	0.003	0.003	0.003
DEE	0.36	0.35	0.32	0.31
FA	0.056	0.083	0.066	0.091
FRIEDMAN	1.36	1.22	1.18	1.19
HO	0.015	0.017	0.017	0.016
HOUSING	32.72	24.82	18.24	14.02
LASER	0.0025	0.0026	0.0023	0.0024
LW	0.026	0.021	0.026	0.033
MORTGAGE	1.22	0.54	0.18	0.032
PLASTIC	3.61	3.27	2.66	2.43
PY	0.13	0.11	0.17	0.22
QUAKE	0.042	0.042	0.045	0.041
SN	0.025	0.027	0.027	0.025
STOCK	6.58	3.40	2.14	1.47
TREASURY	1.83	1.021	0.38	0.11
AVERAGE	**6.34**	**5.56**	**5.04**	**5.10**

4. Conclusions

Author Contributions: V.C. and I.G.T. conducted the experiments, employing several datasets and provided the comparative experiments. 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.

Institutional Review Board 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 conflict 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.