"Feature Clustering-Assisted Feature Selection with Differential Evolution"
Online Supplementary Materials

Peng Wang[a], Bing Xue[a], Jing Liang[b,c], Mengjie Zhang[a]

[a]*School of Engineering and Computer Science, Victoria University of Wellington, Wellington, New Zealand*
[b]*School of Electrical Engineering and Automation, Henan Institute of Technology, Xinxiang, China*
[c]*School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou, China*

1. Introduction

This is the Online Supplementary Materials of "Feature Clustering-Assisted Feature Selection with Differential Evolution".

1.1. Analysis on the Effect of Parameter

For the small constant (τ) in FCNDE, the average training classification accuracy of τ at 0, 0.05, 0.1, 0.15, and 0.2 are reported in Table 1.

For the small constant (τ) in FCNDE, the average training classification accuracy of τ at 0, 0.05, 0.1, 0.15, and 0.2 are reported in Table 1. In Table 1, $\tau = 0$ means only the irrelevant features in a dataset are removed. The five methods show almost the same training accuracy on the SRBCT, Leukemia1, and Leukemia2 datasets. In addition, FCNDE with $\tau = 0.1$ can achieve the best classification performance on the WBCD, Madelon, and TTC3600 datasets. When τ is larger, e.g., 0.15 or 0.2, the performance slightly decreases such as on the AD and TTC3600 datasets. This is because some informative features might be removed with a slightly large τ value. Therefore, this study set τ to 0.1.

1.2. Analysis on the Training Performance

During the evolutionary training process, the calculation of the classification error rate is based on KNN with 5-fold cross-validation on the training set. The average precision and the corresponding standard deviation resulting from the cross-validation

Table 1: Average training classification accuracy (%) of FCNDE using different τ values.

Dataset	$\tau = 0$	$\tau = 0.05$	$\tau = 0.1$	$\tau = 0.15$	$\tau = 0.2$
SPECT	82.64±0.60	**82.81**±0.48≈	82.64±0.77≈	82.29±0.69↓	82.60±0.39≈
WBCD	94.40±0.08	94.39±0.10≈	**94.42**±0.09≈	94.40±0.08≈	94.41±0.06≈
Ionosphere	90.13±0.77	90.14±0.76≈	90.17±0.78≈	**90.24**±0.59≈	89.99±0.92≈
Movement	73.73±0.90	73.59±0.88≈	73.75±0.86≈	73.55±0.98≈	**73.81**±0.84≈
Hillvally	61.81±0.38	**61.88**±0.53≈	61.81±0.43≈	61.82±0.41≈	61.68±0.37≈
Musk1	**98.25**±0.22	98.22±0.14≈	98.17±0.14≈	98.20±0.19≈	98.11±0.14↓
Multiple	98.31±0.13	98.30±0.11≈	98.30±0.12≈	**98.39**±0.14≈	98.32±0.12≈
Arrhythmia	42.47±0.80	42.43±0.82≈	42.48±0.70≈	42.26±0.70≈	**42.54**±0.45≈
Madelon	89.48±0.45	89.36±0.59≈	**89.59**±0.53≈	89.33±0.75≈	89.22±0.47≈
AD	**89.81**±0.89	89.43±0.15≈	88.52±0.0≈	86.53±0.0↓	86.32±0.0↓
SRBCT	**99.99**±0.0	**99.99**±0.0≈	**99.99**±0.0≈	**99.99**±0.0≈	**99.99**±0.0≈
Leukemia1	**99.98**±0.0	**99.98**±0.0≈	**99.98**±0.0≈	**99.98**±0.0≈	**99.98**±0.0≈
TTC3600	79.56±0.44	79.58±0.52≈	**80.43**±0.55↑	79.60±0.31≈	77.13±0.24↓
Prostate	99.34±0.72	99.58±0.70≈	99.56±0.65≈	**99.73**±0.61≈	99.68±0.56≈
Leukemia2	**99.96**±0.0	**99.96**±0.0≈	**99.96**±0.0≈	**99.96**±0.0≈	**99.96**±0.0≈
11Tumor	**86.58**±2.56	85.84±2.37≈	86.00±2.29≈	86.02±2.47≈	85.65±2.13≈

Table 2: Average precision from different methods with the 5-fold cross-validation on the training sets.

Dataset	Full	r3pso	NCDE	LBPADE	BGWO	SBDA	VC-CCPSO	HFS-C-P	FCNDE
SPECT	0.865↓	0.707±0.025↓	0.738±0.022↓	0.790±0.011↓	**0.946**±0.060↑	0.759±0.071↓	0.837±0.009↓	0.774±0.065↓	0.875±0.006
WBCD	0.933↓	0.853±0.014↓	0.892±0.011↓	0.915±0.004↓	0.924±0.013↓	0.905±0.017↓	0.919±0.002↓	0.919±0.014↓	**0.951**±0.001
Ionosphere	0.854↓	0.839±0.004↓	0.864±0.009↓	0.882±0.003↓	0.906±0.020↓	0.898±0.058↓	0.919±0.002↓	0.922±0.007≈	**0.923**±0.006
Movement	0.666↓	0.648±0.003↓	0.661±0.006↓	0.683±0.008↓	**0.961**±0.056↑	0.876±0.146↑	0.735±0.026↓	0.696±0.023↓	0.772±0.013
Hillvally	0.525↓	0.557±0.003↓	0.560±0.001↓	0.572±0.004↓	0.564±0.006↓	0.595±0.020↓	0.581±0.003↓	0.567±0.013↓	**0.621**±0.005
Musk1	0.956↓	0.973±0.002↓	0.978±0.001↓	0.981±0.001≈	0.958±0.003↓	0.953±0.008↓	0.979±0.004↓	0.981±0.002↓	**0.982**±0.001
Multiple	0.975↓	0.974±0.0↓	0.976±0.001↓	0.978±0.001↓	0.978±0.001↓	0.977±0.003↓	0.980±0.001↓	0.980±0.002↓	**0.984**±0.001
Arrhythmia	0.443↓	0.476±0.008↓	0.501±0.005↓	0.522±0.007↓	0.445±0.018↓	0.463±0.029↓	0.499±0.007↓	0.464±0.021↓	**0.634**±0.013
Madelon	0.719↓	0.764±0.007↓	0.785±0.007↓	0.802±0.004↓	0.706±0.015↓	0.624±0.050↓	0.855±0.022↓	0.881±0.003↓	**0.892**±0.005
AD	0.929↓	0.942±0.001↓	0.946±0.001↓	0.950±0.001↓	0.935±0.005↓	0.942±0.004↓	0.934±0.002↓	0.929±0.002↓	**0.954**±0.0
SRBCT	0.827↓	0.870±0.010↓	0.892±0.004↓	0.915±0.008↓	0.835±0.020↓	0.826±0.035↓	0.954±0.012↓	0.887±0.047↓	**1.0**±0.0
Leukemia1	0.788↓	0.861±0.008↓	0.873±0.004↓	0.897±0.009↓	0.790±0.005↓	0.803±0.017↓	0.823±0.008↓	0.876±0.047↓	**1.0**±0.0
TTC3600	0.745↓	0.694±0.006↓	0.692±0.006↓	0.698±0.004↓	0.681±0.007↓	0.696±0.009↓	0.717±0.007↓	0.775±0.007↓	**0.813**±0.006
Prostate	0.747↓	0.820±0.007↓	0.832±0.003↓	0.850±0.006↓	0.751±0.019↓	0.782±0.021↓	0.830±0.004↓	0.843±0.019↓	**0.996**±0.006
Leukemia2	0.915↓	0.935±0.002↓	0.939±0.002↓	0.952±0.005↓	0.912±0.012↓	0.898±0.028↓	0.946±0.002↓	0.934±0.026↓	**1.0**±0.0
11Tumor	0.690↓	0.721±0.008↓	0.721±0.002↓	0.736±0.008↓	0.694±0.013↓	0.689±0.020↓	0.682±0.005↓	0.693±0.019↓	**0.883**±0.023
Sum	16↓	16↓	16↓	1≈,15↓	2↑,14↓	16↓	16↓	1≈,15↓	N/A
Rank	7.3	6.7	5.4	3.5	5.6	6.6	4.3	4.5	1.2

are shown in Table 2. The Wilcoxon test with a significance level of 0.05 is used to test whether there is a statistically significant difference between the proposed FCNDE method and other algorithms. The signs '↑', '↓', and '≈' indicate that the corresponding benchmark algorithm is significantly better than, worse than, and has no significant difference from FCNDE, respectively. The more '↓', the better the proposed FCNDE method. Furthermore, the Friedman test is used to give the relative performance ranking of the algorithms.

As shown in Table 2, the proposed FCNDE method achieves significantly better training precision performance than all the compared baseline methods. Only on the SPECT and Movement datasets, BGWO achieves significantly better precision value than FCNDE. Only on the Ionosphere and Musk1 datasets, FCNDE has a similar precision performance to HFS-C-P and LBPADE, respectively. On 14 out of the 16 datasets, the proposed FCNDE method shows the largest precision value among the nine methods. On three datasets (SRBCT, Leukemia1, and Leukemia2), FCNDE can achieve the top precision value.

1.3. Analysis on the Convergence

The average convergence curves between LBPADE and the proposed FCNDE methods on six datasets are presented in Fig. 1 to further reveal the behavior of the proposed algorithm. On the remaining datasets, similar patterns can be found. The LBPADE method is selected here for comparison with FCNDE since the two methods are ranked the top two in terms of the test classification accuracy according to the results in Table 3 in the main paper.

As shown in Fig. 1, FCNDE achieves significantly lower fitness values than LBPADE on all the six datasets. i.e., the Ionosphere, Hillvally, Arrhythmia, Madelon, SRBCT, and Leukemia2 datasets. The results show that better feature subsets can be produced by the proposed FCNDE method during the training process. Although LBPADE and FCNDE use the same initialization method and the same real-valued encoding, each dimension of an individual in FCNDE corresponds to one feature cluster/group while that in LBPADE corresponds to one feature. Therefore, in the first generation, the average training performance of solutions from the two methods are

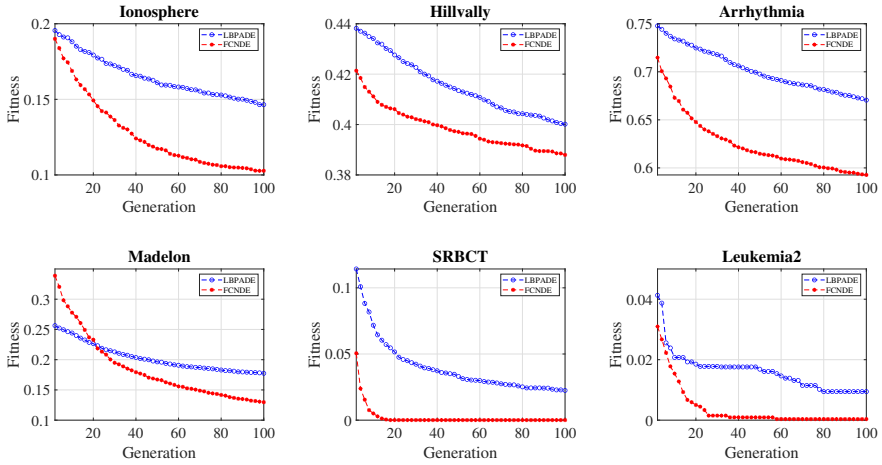

Figure 1: Convergence curves of the two methods on the six datasets.

different.

1.4. Further Results

Table 3 shows the running time of the eight evolutionary feature selection methods. Table 4 gives the average classification accuracy and the number of the selected features of four algorithms, and three of them are variants of the proposed FCNDE method. Fig. 2 shows the unique numbers of feature subsets produced with generations from two methods on the Ionosphere, Arrhythmia, AD, and 11Tumor datasets. The average and the standard deviation of the number of feature subsets including different features but with the same classification accuracy from the four methods, r3pso, NCDE, LBPADE, and FCNDE, are shown in Table 5.

Table 3: Average training time of the eight algorithms (unit: minute).

Dataset	r3pso	NCDE	LBPADE	BGWO	SBDA	VS-CCPSO	HFS-C-P	FCNDE
SPECT	0.9	0.4	0.9	**0.2**	1.2	1.1	1.2	1.2
WBCD	**1.3**	2.0	1.9	1.6	2.2	1.6	1.8	2.8
Ionosphere	1.4	**0.9**	1.9	1.1	2.1	1.8	1.8	2.5
Movement	3.8	**2.9**	7.8	5.0	6.5	4.6	4.8	9.2
Hillvally	**4.7**	25.8	15.8	16.9	13.5	5.4	6.4	26.2
Musk1	**11.0**	128.4	60.5	114.7	71.7	11.8	14.4	64.1
Multiple	18.9	180.8	75.6	372.1	211.2	**18.7**	34.2	106.5
Arrhythmia	**10.4**	11.9	36.3	36.1	27.2	11.7	12.9	68.3
Madelon	42.7	402.8	120.4	1145.9	636.8	**16.0**	18.4	111.5
AD	202.8	1881.1	256.6	1426.2	970.7	97.3	**16.7**	101.1
SRBCT	9.5	**2.8**	103.8	24.0	23.3	10.7	8.7	78.3
Leukemia1	9.8	**6.7**	262.0	39.9	39.3	11.6	9.2	84.4
TTC3600	1080.0	3965.7	829.8	663.8	742.9	313.4	**86.0**	189.6
Prostate	11.8	25.0	421.8	89.8	85.4	14.2	**10.4**	74.7
Leukemia2	11.0	18.8	534.2	72.0	76.6	14.4	**10.0**	110.9
11Tumor	**16.3**	85.0	526.8	234.9	174.5	18.1	45.9	137.8

Table 4: Average accuracy Ac and subset size d^* obtained by the variants of FCNDE.

Algorithms	SPECT		WBCD		Ionosphere		Movement		Hillvally		Musk1		Multiple		Arrhythmia	
	d^*	$Ac(\%)$	d^*	$Ac(\%)$	d^*	$Ac(\%)$	d^*	$Ac(\%)$	d^*	$Ac(\%)$	d^*	$Ac(\%)$	d^*	$Ac(\%)$	d^*	$Ac(\%)$
DE-SU	7.5	76.99	**5.3**	94.13	4.0	83.45	20.3	73.82	**7.7**	54.34	**10.3**	94.37	**45.0**	95.27	21.3	58.16
DE-MIC	6.4	77.53	8.7	**94.15**	4.3	86.79	14.1	74.16	8.1	54.51	12.7	96.40	59.4	96.11	18.5	59.04
NDE-MIC	6.4	**77.67**	7.1	**94.15**	4.3	89.01	**13.3**	74.61	8.2	54.60	14.4	97.18	55.2	96.22	**17.6**	59.43
FCNDE	**6.0**	77.61	8.1	**94.15**	**3.4**	**89.31**	14.2	**74.87**	8.5	**54.78**	14.2	**97.84**	59.8	**96.53**	20.0	**59.75**
Algorithms	Madelon		AD		SRBCT		Leukemia1		TTC3600		Prostate		Leukemia2		11Tumor	
	d^*	$Ac(\%)$	d^*	$Ac(\%)$	d^*	$Ac(\%)$	d^*	$Ac(\%)$	d^*	$Ac(\%)$	d^*	$Ac(\%)$	d^*	$Ac(\%)$	d^*	$Ac(\%)$
DE-SU	**13.8**	87.08	52.1	96.37	85.3	96.33	266.6	76.17	133.4	75.29	181.3	87.25	529.7	89.52	5721.4	75.96
DE-MIC	14.3	87.82	8.1	96.44	84.6	99.01	138.0	81.89	98.0	76.66	144.9	87.57	333.8	91.52	441.7	74.78
NDE-MIC	13.9	88.08	**8.0**	96.43	**62.2**	99.17	**99.1**	83.75	**95.8**	77.12	**123.4**	88.23	**271.6**	92.07	**388.7**	75.92
FCNDE	18.2	**88.25**	**8.0**	**97.45**	89.0	**99.45**	156.5	**84.32**	102.0	**77.28**	165.3	**89.16**	376.2	**92.54**	482.9	**76.43**

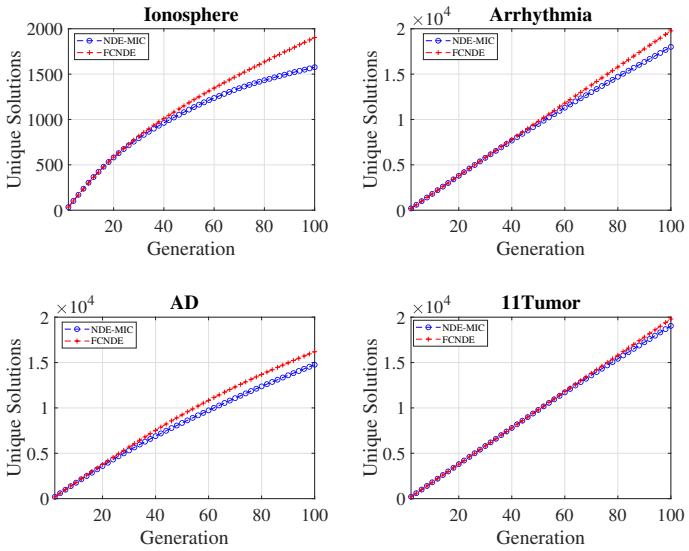

Figure 2: Numbers of unique feature subsets produced with generations from NDE-MIC and FCNDE.

Table 5: Num values of r3pso, NCDE, LBPADE, and FCNDE.

Dataset	r3pso	NCDE	LBPADE	FCNDE
SPECT	**3.4**±1.8	1.0±0	1.1±0.4	2.3±0.4
WBCD	2.5±1.4	1.1±0.3	6.2±2.8	**7.5**±3.1
Ionosphere	**2.5**±1.5	1.0±0	1.1±0.2	2.1±0.8
Movement	**9.6**±3.7	1.0±0	1.0±0	4.3±2.9
Hillvally	1.0±0	1.1±0.2	1.0±0	**3.0**±1.2
Musk1	1.0±0	1.1±0.2	1.0±0	**2.5**±0.9
Multiple	1.1±0.2	1.1±0.2	2.2±1.4	**3.1**±1.4
Arrhythmia	1.0±0	1.0±0	1.0±0	**35.2**±66.4
Madelon	1.0±0	1.0±0	1.0±0	1.0±0
AD	1.0±0	1.0±0	1.3±0.8	**4.6**±0.7
SRBCT	1.2±0.6	1.0±0	3.2±2.9	**63.5**±40.7
Leukemia1	1.1±0.2	1.0±0	8.6±8.3	**91.0**±64.3
TTC3600	1.3±0.6	1.1±0.3	15.2±15.8	**64.3**±52.7
Prostate	1.0±0	1.0±0	16.4±18.5	**151.2**±68.5
Leukemia2	1.0±0	1.1±0.3	**58.9**±66.0	29.3±27.1
11Tumor	1.0±0	1.1±0.2	4.1±4.5	**101.9**±85.8