2-Degree Polynomial Circular Extreme Learning Machine for Classification Problem

Sarutte Atsawaraungsuk
Computer Education Department, Education Faculty, Udonthani Rajabhat University, Udonthani, Thailand.
sarutte@udru.ac.th

***Abstract*—The Circular Extreme Learning Machine (CELM) is the combination of Extreme Learning Machine (ELM) and Circular Back Propagation (CBP). CELM uses the structure same as ELM to get speedy in the training process and gets benefit from CBP architecture by using the 2-degree polynomial to make the better performance. However, the using 2-degree polynomial of CELM can create the decided boundary shape similar as RBF that may be limited the distance, between the center point and the data by the sigmoid activation function. So that, a 2-degree Polynomial Circular Extreme Learning Machine (PCELM) is proposed to tackle this problem. Our experiment showed that PCELM outperformed the original ELM and the traditional CELM with several activation functions. Wilcoxon signed rank test was used to compare statistical differences between PCELM and the CELM that can confirm the PCELM can improve the performance of CELM.**

***Index Terms*—Extreme Learning Machine; Polynomial; Circular Extreme Learning Machine.**

I. INTRODUCTION

Ridella et al. presented the Circular Back Propagation (CBP) that was successfully applied to evaluate the perceived image quality in many application [1-2]. CBP network improves basic input formulation of Multilayer Perceptron (MLP) by adding one dimension, the norm of its data. The additional dimension supports the overall performance of the CBP without affecting the generalization properties with using MLP structure.

Extreme Learning Machine (ELM) [3-5] is very popular with the quickly the learning machine property because ELM generates the random input weights and biases that work on the Single-hidden Layer Feedforward Neural networks (SLFNs). And its efficiency also is satisfiable. Many applied ELMs have developed to enhance the ELM performance for cope several problems.

The CELM [6] is the one applied of ELM that works on the SLFNs structure the same as ELM and uses Circular Back Propagation (CBP) input processing. The CELM gets speedy in the training process from the structure of ELM and fruitful from CBP, the additional dimension that makes the better performance. Moreover, the CELM with activation function has applied in many the of perceived image quality system successfully [7-8].

However, the CELM with the sigmoid activation function may limit the distance, between the center point and the data. Because of CELM can create the decided boundary shape similar as RBF [11-12] (in the range [-2,2] per 1 attribute). But the CELM can calculate hidden layer metric (H) by using the 2-degree polynomial without activation function to unlock the range and the 2-degree polynomial also appropriates with the massive attribute [9-10].

This paper introduces the CELM without activation function, 2-degree Polynomial Circular Extreme Learning Machine (PCELM). The advantage of PCELM is to reduce time in the process of the CELM because PCELM does not use the activation function and PCELM also takes the effect of the 2-degree polynomial to get the accuracy rates better than the traditional CELM with the several activation functions which show in the experimental results.

This article is organized into three parts as follows: the related work presents the background of CBP, ELM, and the proposed PCELM algorithm. The next section is about the experimental results and discussion. And the conclusion is shown in the last.

II. RELATED WORK

A. Circular Back Propagation

Circular Back Propagation (CBP) networks [1] is applied from MLP by adding one dimensional input vector which is the norm of itself, can calculate as follow.

$$\mathbf{r} = b_j + w_j \, \mathrm{x}_i + w_{jD+1} \left\| x_i \right\|^2 \tag{1}$$

where:

$\mathbf{r}$: Stimulus of CBP

The additional term improves the performance of the CBP by not has any affecting to the MLP structure properties.

B. Extreme Learning Machine

Huang et al. [3] proposed ELM. It is very fast learning machines because it is SLFNs that have K hidden nodes. Let N samples that come into the ELM, the sample can be written follow as (x_i, t_i), i=1, 2, …, N where $x_i \in \mathbf{R}^N$ is an training samples and $\mathbf{T} \in \mathbf{R}^C$ is the sample target. In input layer of ELM, the input weights and biases are generated by the random numbers in the range [-1.1] and [0,1], respectively and ELM can be written to the least square form follow as Equation (2).

$$\beta = \mathbf{H}^{\dagger}\mathbf{T} \tag{2}$$

where:

β : Output weights metric which is equal to $\left[\beta_{j1}, \beta_{j2}, \ldots, \beta_{jC}\right]^T$ with j = *1, 2, …, K*

$\mathbf{H}^{\dagger}$: inverse of $\mathbf{H}$ from the Moore-Penrose pseudo inverse

The value of $\mathbf{H}$ is given by Equation (3).

$$\mathbf{H} = h_{ij} = g(w_j * x_i + b_j) \tag{3}$$

where:

w_j : Input weight which is equal to $[w_{j1}, w_{j2}, \dots, w_{jD}]^T$

b_j : Input bias

C. The Proposed Method (PCELM)

The 2-degree Polynomial Circular Extreme Learning Machine (PCELM) is the CELM without using activation function that can calculate hidden layer metric (H) to create the decided boundary shape similar as RBF that shows in Figure 1. The PCELM has the structure as same as ELM but the input vector calculation like the CBP from (1) that can write to new form as:

$$\mathbf{H} = h_{ij} = z_j(\|x_i - c_j\|^2 - b_j),\ j = 1, 2, \dots, K \tag{4}$$

where:

z_j : Equation (5)

c_j : Equation (6)

$$z_j = w_{j,M+1} \tag{5}$$

$$c_j = \left[-\frac{w_{j,1}}{2w_{j,M+1}}, \dots, -\frac{w_{j,M}}{2w_{j,M+1}} \right] \tag{6}$$

$$b_j = \frac{1}{w_{j,M+1}} \left(\sum_{k=1}^{K} \frac{w_{j,M}^2}{4w_{j,M+1}} - br_j \right) \tag{7}$$

where:

w_j : Weight vector which is equal to $[w_{j1}, w_{j2}, \dots, w_{jK}]^T$

$w_{j,M+1}$: Weight vector

b_j : Bias of the hidden layer

br_j : Bias of the hidden layer

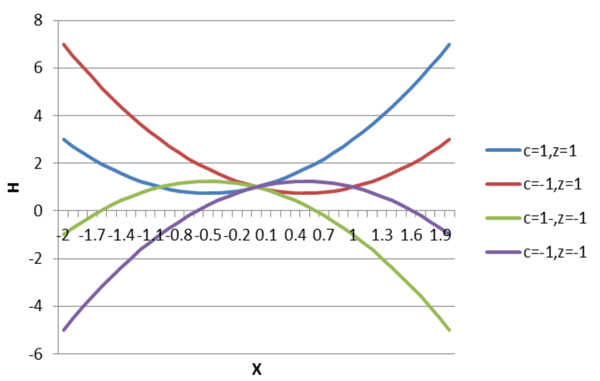

Figure 1: Plotting of hidden layer metric (H) that is calculated by PCELM (1-D) for b=1 using different c and z

From Equation (4) which has the structure like as RBF, weights w are defined the random pattern of X that mean input training data are effect to the center. PCELM is summarized follow as Figure 2.

PCELM Algorithm

Defined

X is the training input data.
K is the number of hidden nodes.

Training Phase

1. Create the PCELM
   a. Centers c_j are calculated to form (6) that w are randomly the patterns of X.
   b. Weights $w_{j,M+1}$ are randomly generated in the range [-1,1].
   c. Widths br_j are generated by the random number in the range [0,1].
2. Calculate **H** from (4).
3. Compute output weight β form (2).

Testing Phase

Calculate **T** to be the result of PCELM.

Figure 2: Algorithm of PCELM

III. Experimental Result

This section is the discussion of the experimental results of the PCELM that used 46 datasets from the UCI repository [13] for analyzing with the compared methods; ELM (sigmoid, RBF), CELM with sigmoid, sine, hard limit, tri bas and RBF. The detail of dataset shows in Table 1. The experiments tested with MATLAB R2014a on environment Core i3-4130 3.40 GHz 4.00 GB. The 10-fold cross was used to validate the input weights and biases parameters that are generated by randomization. To getting the suitable number of hidden nodes was run in the range [1-200]. To evaluate the methods performance many statistical were used to measure the correct answer **T**. Evaluating the methods are consisted: the best accuracy, the win/tie/loss statistical between PCELM and the compared methods in the form meta-metrics [14-15], the mean, median percentage and the ranking.

Table 1
The Detail of All Dataset

Dataset	No. Data	No. Attribute	No. Class
#Lenses	24	4	3
#Iris	150	4	3
#Balance	625	4	3
#Newthyroid	215	5	3
#Liver	345	6	2
#Ecoli	336	7	8
#Pima	768	8	2
#Yeast	1484	8	10
#GlassG2	163	9	2
#Glass	214	9	6
#Breast-w	699	9	2
#Tic-tac-toe	958	9	2
#Vowel	990	10	11
#Wine	178	13	3
#HeartY	270	13	2
#Heart	270	13	2
#Breast	286	15	2
#Zoo	101	16	7
#Vote	435	16	2
#Vehicle	846	18	4
#Hepatitis	155	19	2
#Segment	2310	19	7
#Post-op	90	20	3
#Heart-c	302	22	2
#Primary-tumor	339	23	21
#Labor	27	29	2
#Hypo	3772	29	4

Dataset	No. Data	No. Attribute	No. Class
#Sick	3772	33	2
#Sonar	3772	33	2
#Dermatology	366	34	6
#Ionos	351	34	2
#Satimage	6435	36	6
#Promoters	106	114	2
#Lymph	148	38	4
#Krvskp	3196	38	2
#Waveform	5000	40	3
#Card	690	51	2
#Horse	364	58	3

Dataset	No. Data	No. Attribute	No. Class
#Anneal	898	59	5
#German	1000	61	2
#Optdigits	5320	64	10
#Page-blocks	5473	64	10
#Autos	205	72	6
#Soybean	683	82	19
#Audio	226	93	24
#Promoters	106	114	2
#Gene	3175	120	3

Table 2
The Comparison Performance of PCELM with the Compared Methods

Dataset	ELM (sig)	ELM (RBF)	CELM (sig)	CELM (sine)	CELM (hard limit)	CELM (tri bas)	CELM (RBF)	PCELM
#Lenses	0.7500	0.7667	0.7500	0.8000	0.7000	0.6333	0.7833	0.7667
#Iris	0.9867	0.9800	0.9800	0.9800	0.8800	0.4733	0.9800	0.9800
#Balance	0.9264	0.9233	0.9169	0.9184	0.5903	0.4850	0.9184	0.9184
#Newthyroid	0.9117	0.9160	0.9537	0.9535	0.7636	0.6431	0.9539	0.9582
#Liver	0.7598	0.7533	0.7395	0.7477	0.6004	0.5824	0.7395	0.7396
#Ecoli	0.8634	0.8604	0.8515	0.8513	0.6724	0.3846	0.8485	0.8488
#Pima	0.7682	0.7696	0.7618	0.7630	0.7253	0.6108	0.7683	0.7617
#Yeast	0.6071	0.6071	0.6213	0.6179	0.3430	0.2656	0.6179	0.6179
#GlassG2	0.7107	0.7235	0.7474	0.7301	0.6581	0.5706	0.7415	0.7235
#Glass	0.7141	0.7100	0.7329	0.7286	0.4113	0.3680	0.7195	0.7195
#Breast	0.7447	0.7415	0.7485	0.7451	0.7450	0.6756	0.7416	0.7483
#Tic-tac-toe	0.9916	0.9927	0.9916	0.9927	0.9614	0.6693	0.9906	0.9916
#Vowel	0.9394	0.9495	0.9283	0.9535	0.1798	0.1152	0.9263	0.9505
#Wine	0.9889	1.0000	1.0000	1.0000	0.6908	0.4745	1.0000	1.0000
#HeartY	0.8519	0.8407	0.8556	0.8407	0.8296	0.6333	0.8519	0.8481
#Heart-c	0.8383	0.8048	0.8214	0.8185	0.7654	0.5763	0.8216	0.8316
#Breast-w	0.9685	0.9742	0.9699	0.9685	0.9685	0.7175	0.9699	0.9685
#Zoo	0.9900	0.9900	0.9900	0.9900	0.8018	0.3955	0.9709	0.9900
#Vote	0.9724	0.9655	0.9747	0.9770	0.8688	0.6364	0.9769	0.9770
#Vehicle	0.8712	0.8818	0.8652	0.8641	0.3890	0.3015	0.8581	0.8688
#Hepatitis	0.8904	0.8896	0.9029	0.9033	0.8771	0.7083	0.9092	0.8904
#Segment	0.9541	0.9325	0.9069	0.9597	0.1377	0.1377	0.9056	0.9450
#Post-op	0.7222	0.7444	0.7222	0.7222	0.7333	0.6667	0.7222	0.7333
#Heart	0.8444	0.8074	0.8296	0.8333	0.7889	0.6556	0.8370	0.8370
#Primary-tumor	0.4870	0.4191	0.4838	0.4722	0.2830	0.2475	0.4810	0.4871
#Labor	0.9667	0.9833	0.9833	0.9667	0.8100	0.6900	0.9833	0.9833
#Hypo	0.9290	0.9268	0.9306	0.9311	0.9186	0.9184	0.9313	0.9319
#Sick	0.9491	0.9454	0.9483	0.9523	0.9435	0.9197	0.9507	0.9533
#Sonar	0.8117	0.7983	0.8464	0.8314	0.7017	0.6298	0.8412	0.8414
#Dermatology	0.9892	0.9810	0.9811	0.9782	0.6722	0.3492	0.9837	0.9809
#Ionos	0.9002	0.9032	0.9459	0.9517	0.7125	0.6298	0.9431	0.9431
#Satimage	0.8791	0.8889	0.8903	0.8946	0.5490	0.2614	0.8861	0.8917
Lymph	0.8095	0.7424	0.8243	0.8043	0.6395	0.5143	0.8295	0.8371
#Krvskp	0.9612	0.9099	0.9675	0.9768	0.6133	0.5401	0.9690	0.9775
#Waveform	0.8516	0.8402	0.8610	0.8602	0.4080	0.3580	0.8580	0.8614
#Card	0.8725	0.8551	0.8696	0.8696	0.8101	0.6188	0.8696	0.8754
#Horse	0.6694	0.6450	0.7020	0.6777	0.6534	0.5955	0.7023	0.6720
#Anneal	0.9878	0.9800	0.9811	0.9878	0.7750	0.5602	0.9822	0.9889
#German	0.7540	0.7260	0.7620	0.7610	0.7070	0.6980	0.7620	0.7640

Dataset	ELM (sig)	ELM (RBF)	CELM (sig)	CELM (sine)	CELM (hard limit)	CELM (tri bas)	CELM (RBF)	PCELM
#Optdigits	0.9774	0.9605	0.9715	0.9819	0.2667	0.1368	0.9712	0.9826
#Page-blocks	0.9593	0.9587	0.9605	0.9611	0.9075	0.7913	0.9598	0.9604
#Autos	0.6548	0.6238	0.6771	0.6719	0.4650	0.3176	0.6924	0.6824
#Soybean	0.9488	0.9341	0.9474	0.9502	0.3896	0.1524	0.9459	0.9517
#Audio	0.7656	0.6597	0.8012	0.8138	0.4259	0.2180	0.8184	0.8227
#Promoters	0.8118	0.6536	0.8027	0.8400	0.5745	0.6136	0.8300	0.8964
#Gene	0.8145	0.6041	0.7524	0.8819	0.2825	0.2924	0.7511	0.8954
Mean	0.8547	0.8362	0.8576	0.8625	0.6433	0.5094	0.8586	0.8651
Median	0.8718	0.8711	0.8674	0.8757	0.6954	0.5735	0.8638	0.8911
Number of the highest accuracy	8	8	9	9	0	0	5	19
Win/Tie/Loss	31/3/12	31/6/9	27/4/15	25/6/15	45/1/0	46/0/0	29/7/10	-
Average rank	3.6739	4.4348	3.1739	2.8913	6.8478	7.9348	3.2609	2.3261
SD	2.4656	2.2782	2.3477	2.3352	3.6875	5.5832	2.3091	2.2296

The Table 2 shows that PCELM winning is the highest rank (boldface) in 19 out of 46 datasets, yields the highest mean and the lowest average ranking. The results recommend that PCELM has the performance better than the compared methods, especially the dataset that has the large dimension or attribute such as Audio, Gene, Promoters. The Wilcoxon signed rank test is used to verify the results. Table 3 shows *z* values that dedicate the null-hypothesis of the candidate paired can be rejected or accepted. The rejection case, if z is smaller than -1.96 (plus sign will appear). PCELM has plus sign that means PCELM performs significantly better than all the compared methods except CELM (sine).

Table 3
The Wilcoxon Signed Rank Test for All ELMs

Machines	ELM (sig)	ELM (RBF)	CELM (sig)	CELM (sine)	CELM (hard limit)	CELM (tri bas)	CELM (RBF)	PCELM
ELM(sig)	-	3.0168	-1.0083	**-2.3820**	5.8287	5.9052	-1.1379	**-3.1035**
ELM (RBF)	**-3.0168**	-	**-3.3260**	**-4.0699**	5.8287	5.9052	**-3.1999**	**-4.0055**
CELM (sig)	1.0083	3.3260	-	-0.7192	5.8615	5.9052	0.2721	**-2.3507**
CELM (sine)	2.3820	4.0699	0.7192	-	5.8561	5.9052	0.6738	-1.5726
CELM (hard limit)	**-5.8287**	**-5.8287**	**-5.8615**	**-5.8561**	-	5.7285	**-5.8396**	**-5.8413**
CELM (tri bas)	**-5.9052**	**-5.9052**	**-5.9052**	**-5.9052**	**-5.7285**	-	**-5.9052**	**-5.9052**
CELM (RBF)	1.1379	3.1999	-0.2721	-0.6738	5.8396	5.9052	-	**-2.4212**
PCELM	3.1035	4.0055	2.3507	1.5726	5.8413	5.9052	2.4212	-

IV. Conclusion

The 2-degree Polynomial Circular Extreme Learning Machine (PCELM) can unlock the performance of the CELM by does not use any activation function to create the decided boundary shape similar as RBF. In empirical results show that PCELM can improve the accuracy of CELM, especially the dataset that has large dimension. Furthermore, the PCELM also reduces time in the process of the CELM because PCELM can cut the activation function out of the hidden layer.

Acknowledgment

This research was partially supported by the Faculty of Education and Research and Development Institute, Udonthani Rajabhat University, Thailand.

References

[1] S. Ridella, S. Rovetta and R. Zunino, “Circular backpropagation networks for classification,” IEEE Transactions on Neural Networks, vol. 8, pp. 81-97, 1997.
[2] P. Gastaldo, R. Zounino, I. Heynderickx and E. Vicario, “Circular backpropagation networks for measuring displayed image quality,” ICANN 2002, LNCS 2415,vol. 2415, pp. 1219-1224, 2002.
[3] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: a new learning scheme of feed forward neural networks,” Neural Networks. 2004 IEEE International Joint Conference on Neural Networks, Budapest, Hungary, pp. 985-990, 2004.
[4] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: theory and applications,” Neurocomputing, vol. 70, no. 1-3, pp. 489–501, 2006.
[5] G.-B. Huang, D. H. Wang, and Y. Lan, “Extreme learning machines: A survey,” International Journal of Machine Learning and Cybernetics, vol. 2, no. 2, pp. 107–122, 2011.
[6] S. Decherchi, P. Gastaldo, R. Zunino, E. Cambria, and J. Redi, “Circular-ELM for the reduced-reference assessment of perceived image quality,” Neurocomputing, vol. 102, pp. 78–89, Feb. 2013.

[7] S. Atsawaraungsuk and P. Horata, “Evolutionary circular-ELM for the reduced-reference assessment of perceived image quality,” Lecture Notes in Electrical Engineering, pp. 657–664, 2015.

[8] S. Atsawaraungsuk, P. Horata, K. Sunat, S.Chiewchanwattana, and P. Musigawan, “Evolutionary circular extreme learning machine.” In Computer Science and Engineering Conference (ICSEC 2013), Bangkok, Thailand, pp. 292-297, 2013.

[9] T.M. Cover, “Geometrical and statistical properties of systems of linear inequalities with applications,” IEEE transactions on electronic computers, vol 3, pp.326-334, 1965.

[10] I. Guyon, B. Boser and V. Vapnik, “Automatic capacity tuning of very large VC-dimension classifiers,” Advances in neural information processing systems, pp.147-147, 1993.

[11] G.B. Huang, and C.K. Siew, “Extreme learning machine: RBF network case,” Automation, Robotics and Vision Conference ICARCV 2004 8th, vol. 2, pp. 1029-1036, 2004.

[12] G.B. Huang, and C.K. Siew, Extreme learning machine with randomly assigned RBF kernels. International Journal of Information Technology, vol 11, no 1, pp.16-24, 2005.

[13] M. Lichman , UCI machine learning repository. Irvine, CA: University of California, School of Information and Computer Science. [Online]. Available: http://archive.ics.uci.edu/ml. [Accessed: 4-Jan-2016].

[14] A. Stefani and M. Xenos, “Meta-metric evaluation of E-commercerelated metrics,” Electronic Notes in Theoretical Computer Science, vol. 233, pp. 59–72, 2009.

[15] P. Horata, S. Chiewchanwattana, and K. Sunat. “Robust extreme learning machine.” Neurocomputing, vol. 102, pp. 31-44, 2013.