APPLICATION OF REGULARIZED ONLINE SEQUENTIAL LEARNING FOR HEMATOCRIT ESTIMATION

HIEU TRUNG HUYNH[1] AND YONGGWAN WON[2]

[1]*Faculty of Information Technology, Industrial university of Ho Chi Minh city, Viet Nam*
[2]*Department of Computer Engineering, Chonnam National University, Gwangju 500-757, Korea*
hthieu@iuh.edu.vn

Abstract. Hematocrit (HCT) is expressed as the percentage of red blood cells in the whole blood, it is one of the most highly affecting factors which influences the glucose measurement by using handheld device. In this paper, we present an approach for applying the regularized online sequential learning to hematocrit estimation. The input is the transduced current curve which is produced by the chemical reaction during glucose measurement. The experimental results shown that the proposed approach is promising.
Keywords. hematocrit; neural network; online training; extreme learning machine; handheld device.
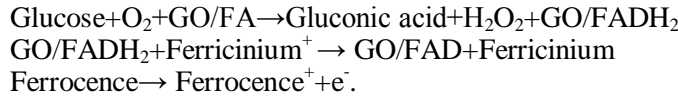
1. INTRODUCTION

The neural network is widely applied in several applications [1-4] due to its abilities to solve problems which are difficult to handle by using traditional approaches and to approximate complex nonlinear mappings directly from input patterns. Several network architectures have been developed, however it was shown that the single hidden layer feedforward neural networks (SLFN) can approximate any function if the activation function is chosen properly. Hence, in this study, we have investigated in the SLFN for biomedical processing. Several training algorithms have been developed for SLFNs, in which one of the effective ones is extreme learning machine (ELM) [5, 6]. This algorithm can obtain good performance with higher learning speed in many applications. Besides batch learning types, sequential learning algorithms are preferred for neural networks in many applications, they do not require the fully available training set and do not require retraining whether a new training data received. In this paper, we propose an approach that applies the regularized online sequential learning algorithm for hematocrit estimation.

Hematocrit (HCT) is one of useful clinical indicators in surgical procedures and hemodialysis, and anemia [7-9]. It is also a factor highly affecting the accuracy of glucose measurements [10-12]. The glucose values are trended to underestimation at higher hematocrit levels and overestimation at lower hematocrit levels. Hence, one of approaches to improve the accuracy of glucose measurements in the handheld devices is to reduce the effects of HCT [13]. The hematocrit can be measured directly by centrifugation in a small laboratory. Most commonly, it is measured indirectly by an automated blood cell counter. It also can be estimated by dielectric spectroscopy [14] or some different techniques. As most of the above approaches require individual devices or are quite complicated, the proposed methods for estimating hematocrit by using the glucose biosensors which can be used to correct the glucose measurements and integrated into the handheld meters for glucose measurement [15-16]. In this study, we present an application of the regularized online sequential extreme learning machine for hematocrit estimation. The rest of this paper is organized as follow. Section 3 presents the proposed approach for estimating hematocrit. The experimental results and analysis are shown in section 3. Finally, we make the conclusion in section 4.

2. THE REGULARIZED ONLINE SEQUENTIAL LEARNING ALGORITHM FOR HEMATOCRIT ESTIMATION

2.1 Transduced current curves

The online sequential learning for estimating hematocrit response has the input from transduced current curves. These curves are produced by the chemical reaction between the enzyme coated on the biosensor test strips and blood. One of enzymes commonly used in biosensors to detect the glucose levels is the glucose oxidase (GOD) which is used to catalyze the oxidation of glucose by oxygen to produce gluconic acid and hydrogen peroxide.

$$\text{Glucose} + O_2 + GO/FA \rightarrow \text{Gluconic acid} + H_2O_2 + GO/FADH_2$$
$$GO/FADH_2 + \text{Ferricinium}^+ \rightarrow GO/FAD + \text{Ferricinium}$$
$$\text{Ferrocence} \rightarrow \text{Ferrocence}^+ + e^-.$$

The reduced form of the enzyme ($GO/FADH_2$) is oxidized to its original state by an electron mediator (ferrocence). The active electrode then oxidizes the resulting reduced mediator to produce the transduced anodic current. The transduced anodic current curve obtained in the first 14 seconds is represented in Fig. 1 [17]. It was shown that the first eight seconds do not contain the information of hematocrit; it may be an incubation time for waiting the enzyme reaction to be activated. In our study, we concentrate on the second part of the current curve during the next six seconds. In the period of the next six seconds, the anodic current curve is sampled at a frequency of 10Hz to produce current points. The vector of $d=59$ current points sampled from the second part of the j-th current curve can be denoted as $\mathbf{x}_j=[x_{j1}, x_{j2}, ..., x_{j59}]$. This vector is used as the input values of the neural network for estimating hematocrit.
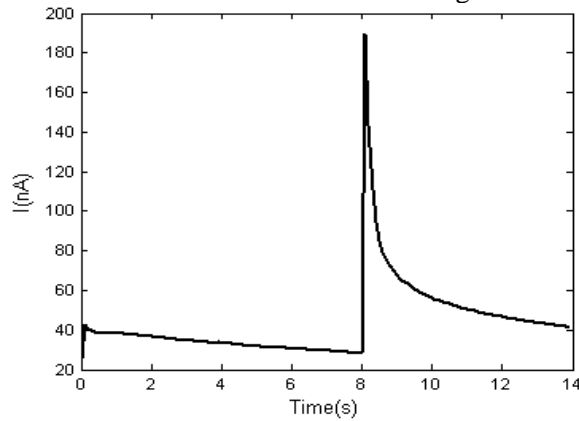

Figure 1. Anodic Current Curve

2.2 Neural networks trained by online training algorithms for hematocrit estimation

The architecture neural network using in this study is single hidden layer feedforward neural network (SLFN) which can approximate any function if the number of hidden nodes and the activation function are chosen properly. The typical architecture of SLFN is shown in Fig. 2, which includes d input nodes, N hidden nodes and C output nodes.
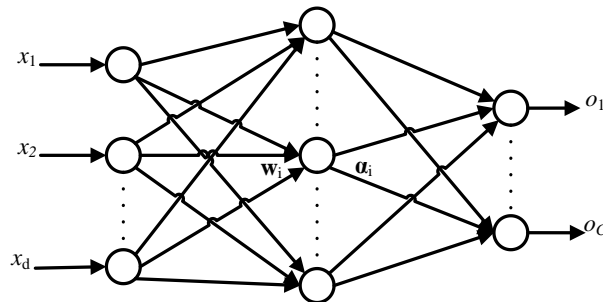

Figure 2. The architecture of SLFN

Let $f(\cdot)$ be the activation function of hidden units. Mathematically, the SLFNs can be modeled as:

$$\mathbf{o} = \sum_{i=1}^{N} \boldsymbol{\alpha}_i f(\mathbf{w}_i \cdot \mathbf{x} + b_i), \ \mathbf{x} \in \mathbb{R}^d, \tag{1}$$

where $\mathbf{o}$ is the output vector, $\mathbf{w}_i=[w_{i1}, w_{i2}, ..., w_{iN}]$ is the input weight vector connecting from the input units to the i-th hidden unit, $\boldsymbol{\alpha}_i$ is the weight vector connecting from the i-th hidden unit to the output units, and b_i is the threshold of the i-th hidden unit, $\mathbf{w}_i \cdot \mathbf{x} = <\mathbf{w}_i, \mathbf{x}>$ is the inner product of $\mathbf{w}_i$ and $\mathbf{x}$. One of big problems in neural networks is training.

Given n training patterns $(\mathbf{x}_j, \mathbf{t}_j)$, $j=1, 2, \ldots, n$, where $\mathbf{x}_j=[x_{j1}\ x_{j2}\ \ldots\ x_{jd}]^T$ and $\mathbf{t}_j=[t_{j1}\ t_{j2}\ \ldots\ t_{jC}]^T$ are the j-th input pattern and its target, respectively. The main goal of training process is to determine the network weights $\mathbf{w}_i$, $\boldsymbol{\alpha}_i$, and biases b_i that minimize the error function defined by

$$E = \sum_{j=1}^{n} \left(\mathbf{o}_j - \mathbf{t}_j \right)^2 , \tag{2}$$

where $\mathbf{o}_j$ is the output vector corresponding to the j-th input pattern. Traditionally, this task is performed based on the gradient descent, in which the network weights $\mathbf{g}$ (consisting of $\mathbf{w}$, $\boldsymbol{\alpha}$ and b) are updated iteratively by:

$$\mathbf{g}_k = \mathbf{g}_{k-1} - \eta \frac{\partial E}{\partial \mathbf{g}} , \tag{3}$$

where η is the learning rate. One of the most popular training algorithms based on gradient descent is backpropagation, in which the network weights are updated from the output layer to the input layer. This algorithm has some problems such as local minima, overtraining, learning rate, etc. There are some improvements for neural networks developed by different research groups. However, up to now, most training algorithms based on gradient descent are still slow due to iterative processes [18-20].

One of effective training algorithms which can overcome some problems in the gradient descent based ones is extreme learning machine (ELM). Let $\mathbf{H}$ be the hidden-layer-output matrix of SLFN which was defined as [5, 6]:

$$\mathbf{H}=\begin{bmatrix} f(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & f(\mathbf{w}_N \cdot \mathbf{x}_1 + b_N) \\ \vdots & \ddots & \vdots \\ f(\mathbf{w}_1 \cdot \mathbf{x}_n + b_1) & \cdots & f(\mathbf{w}_N \cdot \mathbf{x}_n + b_N) \end{bmatrix}, \tag{4}$$

The main goal in ELM is to determine the network weights based on the linear model defined by

$$\mathbf{HA}=\mathbf{T}, \tag{5}$$

where $\mathbf{T}= [\mathbf{t}_1\ \mathbf{t}_2\ \ldots\ \mathbf{t}_n]^T$, $\mathbf{A}=[\boldsymbol{\alpha}_1\ \boldsymbol{\alpha}_2\ \ldots\ \boldsymbol{\alpha}_N]^T$. In the ELM, the input weights and biases of hidden units are randomly assigned, and the output weights are determined by

$$\hat{\mathbf{A}}=\mathbf{H}^{\dagger}\mathbf{T}, \tag{6}$$

where $\mathbf{H}^{\dagger}$ is the pseudo-inverse of $\mathbf{H}$.

When the training data is very large or not available fully, the online training approaches should be addressed. An online training method based on the ELM called sequential extreme learning machine (OS-ELM) was proposed by Liang *et al.* [21]. The OS-ELM supposes that $\mathbf{H}^T\mathbf{H}$ is nonsingular and pseudo-inverse of $\mathbf{H}$ is given by

$$\mathbf{H}^{\dagger}=(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T. \tag{7}$$

From above assumptions, the output weights are updated by following rules:

$$\mathbf{A}_k = \mathbf{A}_{k-1} + \mathbf{L}_k^{-1}\mathbf{H}_k^T(\mathbf{T}_k - \mathbf{H}_k\mathbf{A}_{k-1}) , \tag{8}$$

$$\mathbf{L}_k = \mathbf{L}_{k-1} + \mathbf{H}_k^T\mathbf{H}_k \tag{9}$$

where $\mathbf{T}_k = [\mathbf{t}_{\sum_{i=0}^{k-1} n_i +1}\ \mathbf{t}_{\sum_{i=0}^{k-1} n_i +2}\ \ldots\ \mathbf{t}_{\sum_{i=0}^{k} n_i}]^T$ and $\mathbf{H}_k = [\mathbf{h}_{\sum_{i=0}^{k-1} n_i +1}\ \mathbf{h}_{\sum_{i=0}^{k-1} n_i +2}\ \ldots\ \mathbf{h}_{\sum_{i=0}^{k} n_i}]^T$. The initialization of $\mathbf{A}_k$ corresponding to an initial training set $\mathbf{S}_0=\{(\mathbf{x}_j, \mathbf{t}_j)\ |\ j=1,\ldots, n_0\}$ is given by

$$\mathbf{A}_0= \mathbf{L}_0^{-1}\mathbf{H}_0^T\mathbf{T}_0 , \tag{10}$$

where $\mathbf{L}_0 = \mathbf{H}_0^T\mathbf{H}_0$, $\mathbf{T}_0= [\mathbf{t}_1\ \mathbf{t}_2\ \ldots\ \mathbf{t}_{n0}]^T$, and $\mathbf{H}_0= [\mathbf{h}_1\ \mathbf{h}_2\ \ldots\ \mathbf{h}_{n0}]^T$. In summation, the OS-ELM algorithm as follows:

1) **Initialization**:

For the initial training subset $S_0 = \{(\mathbf{x}_j, \mathbf{t}_j) \mid j=1,\ldots,n_0\}$,

- Assign random values for $\mathbf{w}$'s and b's.
- Calculate hidden layer output matrix $\mathbf{H}_0$.
- Determine $\mathbf{L}_0$ and then $\mathbf{A}_0$ using by using Eq. 10.

2) **Updating weight:** For the arriving training subset $\mathbf{S}_k = \{(\mathbf{x}_j, \mathbf{t}_j) \mid j = \sum_{i=0}^{k-1} n_i + 1, \ldots, \sum_{i=0}^{k} n_i\}$,

- Determine $\mathbf{H}_k$.
- Determine $\mathbf{L}_k$ by Eq. 9.
- Update the output weights $\mathbf{A}_k$ by Eq. 8.

In the first step of algorithm (*initialization*) the input weights and biases are assigned by random values; then the output weight matrix $\mathbf{A}_0$ is computed. Following the initialization step, the updating process is performed, in which the output weights are updated for each arriving data of one-by-one or chunk-by-chunk.

In the real applications, the collected data are often included noise. Hence, the risk minimization as shown in (2) may lead to a poor generalization. One of approaches which can overcome this problem is to optimize the norm of output weight vector. The solution for $\mathbf{A}$ of Eq. 5 can be replaced by seeking $\mathbf{A}$ that minimizes

$$\|\mathbf{HA} - \mathbf{T}\|^2 + \lambda \|\mathbf{A}\|^2, \tag{11}$$

where $\|\cdot\|$ is Euclidean norm and λ is a positive constant. The solution for $\mathbf{A}$ from Eq. 11 is given by

$$\hat{\mathbf{A}} = (\mathbf{H}^T\mathbf{H} + \lambda\mathbf{I})^{-1}\mathbf{H}^T\mathbf{T}. \tag{12}$$

The learning rules for online sequential learning process were given by Hieu TH *et al* [22]. For an initial training set $\mathbf{S}_0 = \{(\mathbf{x}_j, \mathbf{t}_j) \mid j=1,\ldots,n_0\}$, the output weights are initialized by

$$\mathbf{A}_0 = \mathbf{L}_0^{-1}\mathbf{H}_0^T\mathbf{T}_0, \tag{13}$$

where $\mathbf{L}_0 = \mathbf{H}_0^T\mathbf{H}_0 + \lambda\mathbf{I}$, $\mathbf{T}_0 = [\mathbf{t}_1 \ \mathbf{t}_2 \ \ldots \ \mathbf{t}_{n0}]^T$, and $\mathbf{H}_0 = [\mathbf{h}_1 \ \mathbf{h}_2 \ \ldots \ \mathbf{h}_{n0}]^T$. In the updating phase, the output weights are updated by :

$$\mathbf{U}_k = \mathbf{U}_{k-1} - \mathbf{U}_{k-1}\mathbf{H}_k^T(\mathbf{I} + \mathbf{H}_k\mathbf{U}_{k-1}\mathbf{H}_k^T)^{-1}\mathbf{H}_k\mathbf{U}_{k-1} \tag{14}$$

$$\mathbf{A}_k = \mathbf{A}_{k-1} + \mathbf{U}_k\mathbf{H}_k^T(\mathbf{T}_k - \mathbf{H}_k\mathbf{A}_{k-1}). \tag{15}$$

where

$$\begin{aligned}\mathbf{U}_0 &= (\mathbf{H}_0^T\mathbf{H}_0 + \lambda\mathbf{I})^{-1} \\ &= \frac{1}{\lambda}\mathbf{I} - \frac{1}{\lambda}\mathbf{H}_0^T(\lambda\mathbf{I} + \mathbf{H}_0\mathbf{H}_0^T)^{-1}\mathbf{H}_0\end{aligned} \tag{16}$$

3. RESULTS AND DISCUSSIONS

In this study, we evaluate the performance on the dataset which was obtained from 199 blood samples. These samples were obtained from randomly selected volunteers, every sample is divided into two parts, the first part is to determine the anodic current curves, and the second part is to determine the accurate hematocrit using the centrifugation method. From the second part of curve, which is after the incubation time, fifty-nine current points are sampled at a frequency of 10Hz. There are 60 features for every sample, in which 59 features can be considered as input features. The hematocrit values collected from centrifugation method have the distribution as shown in Fig. 3, in which the mean is 36.02 and the deviation is 6.39. The dataset was divided into two subsets, in which the forty percent of dataset is used for training and the sixty percent is used for blind testing. In our experiment, the neural networks were trained by the OS-ELM

[23] our proposed method and offline ELM. The number of hidden units was 12 for ELM and online training algorithms.

The average result of fifty trials with the whole current curve is shown in Table 1. The root mean square error (RMSE) was computed by

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(o_j - t_j)^2} \qquad (11)$$

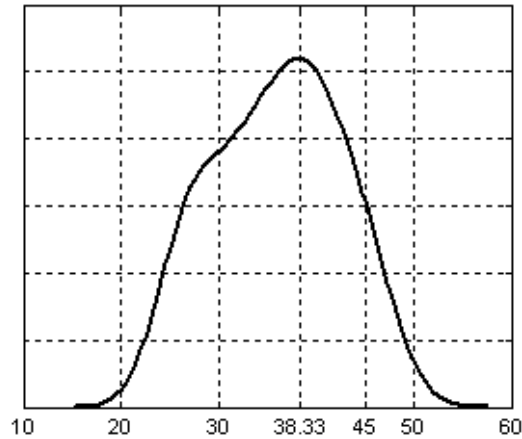where o_j is the estimated value and t_j is the reference value.

Figure 3. Distribution of collected hematocrit

Table 1. Comparison with reference hematocrit measurements using centrifugation (whole current curve).

Method	Training		Testing		# nodes
	RMSE	Std	RMSE	Std	
ELM (offline)	3.67	0.34	4.49	0.51	12
OS-ELM [23]	3.69	0.26	4.37	0.37	12
Proposed approach	3.65	0.28	4.18	0.35	12

From the Table 1 we can see that the accuracy of the proposed method corresponding to the testing set is 4.18 which is compatible to that of the offline other online training methods for the same number of hidden nodes. Note that, for the online training method, the devices can be still trained with new samples during the using process which can expect to improve the performance further.

4.   CONCLUSION

In this paper, we presented a method for hematocrit estimation using the online sequential method based on extreme learning machine. The transduced current changing curves produced by reactions of glucose oxidase in the electrochemical biosensors was used as the input features. The experimental results shown that the online training method is compatible to the offline training methods, but note that the accuracy of devices can be still improve during the using process. This result can be contributed to reduce the hematocrit dependency in measurement of glucose value by electrochemical biosensors.

REFERENCES

[1]   P. J. G. Lisboa, E. C. Ifeachor, and P. S. Szczepaniak: Artificial neural networks in Biomedicine, Springer-Verlag London Berlin Heidelberg (2000)

[2]   R. N. G. Naguib, G. V. Sherbet: Artificial Neural Networks in Cancer Diagnosis, Prognosis, and Patient Management. CRC Press, Washington D.C. (2001)

[3]   Hieu Trung Huynh, Jungja Kim, and Yonggwan Won: Performance comparison of SLFN training algorithms for DNA microarray classification. Advances in Experimental Medicine and Biology, volume 696, pp 135-143 (2011)

[4]   Hieu Trung Huynh, Jungja Kim, and Yonggwan Won: Classification Study on DNA Microarray with Feedforward Neural network Trained by Singular Value Decomposition. International Journal of Bioscience and Biotechnology, vol. 1, no. 1, pp. 17-24 (2009)

[5]   G.-B. Huang, Q.-Y. Zhu, C.-K. Siew: Extreme learning machine: A New Learning Scheme for Feedforward Neural Networks. Proc. of Int'l Joint Conf. on Neural Networks (July 2004)

[6]   G.-B. Huang, Q.-Y. Zhu, C.-K. Siew: Extreme learning machine: Theory and applications. Neurocomputing, vol. 70, pp. 489-501 (2006)

[7]   J. Z. Ma, J. Ebben, H. Xia, and A. J. Collins: Hematocrit level and associated mortality in hemodialysis patients. J. Amer. Soc. Nephrol., vol. 10, pp. 610-619 (1999)

[8]   A. Aris, J. M. Padro, J. O. Bonnin, and J. M. Caralps: Prediction of hematocrit changes in open-heart surgery without blood transfusion. J. Cardiovasc. Surg., vol. 25, no. 6, pp. 545-548 (Nov-Dec. 1984)

[9]   I. F. Bejarano, J. E. Dipierri, E. L. Alfaro, C. Tortora, T. García, M. C. Buys: Hematocrit values and prevalence of anemia in schoolchildren of Jujuy. Medicina, vol. 63, no. 4, pp. 288-292 (2003)

[10] R. F. Louie, Z. Tang, D. V. Sutton, J. H. Lee, G. J. Kost: Point of Care Glucose Testing: effects of critical variables, influence of reference instruments, and a modular glucose meter design. Arch Pathol Lab Med, vol. 124, pp. 257-266 (2000)

[11] Z. Tang, J. H. Lee, R. F. Louie, G. J. Kost, D. V. Sutton: Effects of Different Hematocrit Levels on Glucose Measurements with HandHeld Meters for Point of Care Testing. Arch Pathol Lab Med, vol. 124, pp. 1135-1140 (2000)

[12] E. S. Kilpatrick, A. G. Rumley, H. Myin: The effect of variations in hematocrit, mean cell volume and red blood count on reagent strip tests for glucose. Ann Clin Biochem, vol. 30, pp. 485-487 (1993)

[13] Hieu Trung Huynh, Ho Dac Quan, and Yonggwan Won, "Accuracy Improvement for Glucose Measurement in Handheld Devices by using Neural Networks", Future Data and Security Engineering, FDSE2017, Lecture Notes in Computer Science (LNCS 10640), pp. 299-308 (2017).

[14] E. F. Treo, C. J. Felice, M. C. Tirado, M. E. Valentinuzzi, and D. O. Cervantes: Hematocrit Measurement by Dielectric Spectroscopy. IEEE Transactions on Biomedical Engineering, vol. 25, no. 1, pp. 124-127 (2005)

[15] Hieu Trung Huynh, Yonggwan Won and Jung-ja Kim: Neural Networks for the Estimation of Hematocrit from Transduced Current Curves. The 2008 IEEE Int'l conference on Networking, Sensing and Control, pp.1517-1520 (2008)

[16] Hieu Trung Huynh and Yonggwan Won, "Hematocrit Estimation on Online Trained Neural Network", 2014 International Symposium on Information Technology Convergence (ISITC 2014), pp. 130 – 132 (2014).

[17] Hieu Trung Huynh, Yonggwan Won and Jung-ja Kim: Neural Networks for the Estimation of Hematocrit from Transduced Current Curves. The 2008 IEEE Int'l conference on Networking, Sensing and Control, pp.1517-1520 (2008).

[18] Syed Muhammad Aqil Burney, Tahseen Ahmed Jilani, and Cemal Ardil: A Comparison of First and Second Order Training Algorithms for Artificial Neural Networks. International Journal of Computational Intelligence, vol. 1, pp.218-224 (2004)

[19] L. S. Nghia, J. Sjöberg and M. Viberg: Adaptive neural nets filter using a recursive levenberg-marquardt search direction. Proc. Asilomar Conf. Signals, Syst., Comput., vol. 1-4, pp.697-701 (Nov. 1998).

[20] V. S. Asirvadam, S. F. McLoone, and G.W. Irwin: Parallel and separable recursive Levenberg-Marquardt training algorithm. Proc. 12th IEEE Workshop Neural Net. Signal Process., no. 4-6, pp. 129-138 (Sep. 2002)

[21] N.-Y. Liang, G.-H. Huang, P. Saratchandran, and N. Sundararajan: A fast and accurate online sequential learning algorithm for feedforward networks. IEEE Trans. on Neural Networks, vol. 17, pp. 1411-1423 (2006)

[22] Hieu Trung Huynh and Yonggwan Won, "Regularized online sequential learning algorithm for single-hidden layer
feedforward neural networks", Pattern Recognition Letters 32, pp. 1930-1935 (2011)

[23] Hieu Trung Huynh, Yonggwan Won, and Jungja Kim, "Hematocrit estimation using online sequential extreme learning machine", Bio-Medial Materials and Engineering 26, pp. S2025-S2032 (2015).