Research Article

Risk Stratification with Extreme Learning Machine: A Retrospective Study on Emergency Department Patients

Nan Liu,[1] Jiuwen Cao,[2] Zhi Xiong Koh,[1] Pin Pin Pek,[1] and Marcus Eng Hock Ong[1,3]

[1] Department of Emergency Medicine, Singapore General Hospital, Singapore 169608
[2] Institute of Information and Control, Hangzhou Dianzi University, Zhejiang 310018, China
[3] Health Services and Systems Research, Duke-NUS Graduate Medical School, Singapore 169857

Correspondence should be addressed to Nan Liu; liu.nan@sgh.com.sg

This paper presents a novel risk stratification method using extreme learning machine (ELM). ELM was integrated into a scoring system to identify the risk of cardiac arrest in emergency department (ED) patients. The experiments were conducted on a cohort of 1025 critically ill patients presented to the ED of a tertiary hospital. ELM and voting based ELM (V-ELM) were evaluated. To enhance the prediction performance, we proposed a selective V-ELM (SV-ELM) algorithm. The results showed that ELM based scoring methods outperformed support vector machine (SVM) based scoring method in the receiver operation characteristic analysis.

1. Introduction

In the emergency department (ED), the process of triage enables rapid screening of patients to determine severity and assign proper treatment. Most risk stratification systems are based on clinical judgment and traditional vital signs such as blood pressure and heart rate. However, vital signs alone may not be sufficient for accurate risk assessment. Machine learning has been used to design an Euclidean distance based scoring system (DIST) [1] and showed its advantage over statistical risk scores [2, 3]. Motivated by this encouraging discovery, we aim to derive the DIST score using the extreme learning machine (ELM) [4], one of the latest advancements in machine learning community.

Since ELM was proposed for single-layer feedforward neural networks (SLFNs) [4], it has received many interests on improvement and enhancement at the algorithm level [5–9]. With the development of ELM based methodologies, ELM and its variants have been applied to various applications [10–13]. ELM has been widely applied for classification problems in biomedical signal processing such as electroencephalography (EEG) [14, 15] and electrocardiogram (ECG) [16, 17]. Bioinformatics is another popular area of its application [18, 19]. Both theoretical and experimental studies have shown

the evidence that ELM methods are comparable to SVM but benefit from low computational complexity [20].

To our knowledge, ELM has yet been applied for risk stratification in biomedical applications. In this paper, ELM is employed to enhance the DIST scoring system for effective patient outcome prediction, where heart rate variability (HRV) parameters and vital signs are used as predictors. The rest of the paper is organized as follows. Section 2 elaborates the DIST risk stratification system for patient outcome prediction. Section 3 starts with the introduction of basic ELM algorithm and then presents the voting based ELM (V-ELM) and our proposed selective voting based ELM (SV-ELM). Section 4 describes the experimental setting and data collection, as well as the experimental results. Section 5 concludes the study.

2. Risk Stratification System

We previously developed an Euclidean distance based scoring system (DIST) [1] for risk stratifying critically ill patients presented to the ED. The purpose of the scoring method is to stratify patients into different risk levels so that a proper triage is able to be done. Triage is the process of determining the priority of patients' treatments based on the severity of
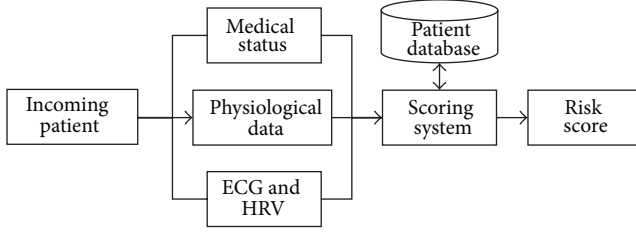
FIGURE 1: The distance based scoring system (DIST) triage architecture. Firstly, the selected variables are converted into interval $[-1, 1]$ with min-max normalization. Secondly, cluster centers for both positive samples (patients with cardiac arrest within 72 h) and negative samples (patients without cardiac arrest within 72 h) are calculated based on Euclidean distance, and an initial score for a testing sample is derived by measuring distances between the testing sample and two cluster centers. Lastly, the classification algorithm is implemented to fine-tune the risk score.

their condition. Risk stratification of patients aims to facilitate the allocation of scarce ED resources, that is, personnel, equipment, and beds.

The DIST risk stratification system is illustrated in Figure 1. It is designed to be applied onto an incoming patient presented to the ED. The DIST system is intended to be a clinical tool to assess incoming patients and to provide a risk score as an output. A low score indicates that the patient is not in critical condition, while a high score indicates the imminent possibility of cardiac arrest.

As an overview, the DIST system utilizes physiological and cardiac data measurements, compiled with medical status information, and processes such inputs within an intelligent machine learning scoring algorithm which compares the present input to correlated past patient diagnoses, in order to provide an insightful risk score as to the risk of cardiac arrest in the patient. A computer interface is provided to an ED nurse to register the incoming patient and to enter pertinent information relating to the medical status input of the patient. The medical status is thereafter transmitted and logged into the triage system under an identifier for the patient. The system is also capable of polling the relevant patient data, retrieving the information required, and propagating the information into the present triage assessment.

Two types of features used in the DIST triage system are heart rate variability (HRV) measures and vital signs. HRV is defined as the variation in the time interval between successive heart beats. Following the widely used HRV analysis standard [21], two categories of HRV (time domain and frequency domain) are calculated. Vital signs are physiological measures of the patient. Vital sign data may be defined as clinical measurements that indicate the state of patient's essential body functions. For example, they may refer to heart rate, respiratory rate, and blood pressure reading.

3. Risk Stratification with ELM

The DIST risk stratification system employs support vector machine (SVM) [22] as the core of its prediction module. Compared with SVM, ELM shows several advantages that

have been discussed [20, 23]. In this study, we aim to apply ELM and the voting base ELM (V-ELM) [24] to evaluate the performance of ELM for risk stratification. Furthermore, we propose a selective V-ELM (SV-ELM) algorithm to enhance the prediction performance.

3.1. Basic ELM. As a fast learning algorithm for single-layer feed-forward network (SLFN), ELM [4] randomly selects weights and biases for hidden nodes and analytically determines the output weights by finding least square solution. Suppose that there are N samples in the training set

$$L = \left\{ \left(\mathbf{x}_j, \mathbf{t}_j \right) \mid \mathbf{x}_j \in \mathbf{R}^p, \ \mathbf{t}_j \in \mathbf{R}^q, \ j = 1, 2, \ldots, N \right\}, \quad (1)$$

where $\mathbf{x}_j$ is a $p \times 1$ input vector and $\mathbf{t}_j$ is a $q \times 1$ target vector. Given that $\mathbf{w}_i$ is p-dimensional weight vector connecting ith hidden node and input neurons and $g(x)$ is the activation function, an SLFN with $\widetilde{N}$ hidden nodes is formulated as

$$f_{\widetilde{N}} \left(\mathbf{x}_j \right) = \sum_{i=1}^{\widetilde{N}} \boldsymbol{\beta}_i g \left(\mathbf{w}_i \cdot \mathbf{x}_j + b_i \right) = \mathbf{t}_j, \quad j = 1, 2, \ldots, N. \quad (2)$$

Then a compact format of (2) can be written as

$$\mathbf{H}\widehat{\boldsymbol{\beta}} = \mathbf{T}, \quad (3)$$

where $\mathbf{H}(\mathbf{w}_1, \ldots, \mathbf{w}_{\widetilde{N}}, b_1, \ldots, b_{\widetilde{N}}, \mathbf{x}_1, \ldots, \mathbf{x}_N)$ is hidden layer output matrix of the network; $h_{ji} = g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i)$ is the output of ith hidden neuron with respect to $\mathbf{x}_j$, $i = 1, 2, \ldots, \widetilde{N}$ and $j = 1, 2, \ldots, N$; $\widehat{\boldsymbol{\beta}} = [\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_{\widetilde{N}}]^{\mathrm{T}}$ and $\mathbf{T} = [\mathbf{t}_1, \ldots, \mathbf{t}_N]^{\mathrm{T}}$ are output weight matrix and target matrix, respectively. To obtain small nonzero training error, Huang et al. [4] proposed randomly assigning values for parameters $\mathbf{w}_i$ and b_i, and thus the system becomes linear so that the output weights can be estimated as $\boldsymbol{\beta} = \mathbf{H}^{\dagger}\mathbf{T}$, where $\mathbf{H}^{\dagger}$ is the Moore-Penrose generalized inverse [25] of the hidden layer output matrix $\mathbf{H}$.

3.2. V-ELM and SV-ELM. The V-ELM [24] method was proposed to improve the classification performance of ELM. Its assumption is that the hidden nodes in basic ELM are randomized and remain unchanged during the training phase, which increases the possibility of misclassification for some samples near the decision boundary. V-ELM utilizes a majority voting mechanism to combine an ensemble of individual basic ELM based decisions. This strategy is reported to well address the misclassification problems on some borderline samples [19, 24]. Suppose that there are K independent networks trained in V-ELM. For each testing sample $\mathbf{x}^{\text{test}}$, K prediction results can be obtained based on these independent ELMs. A corresponding vector $\mathbf{S}_{K,\mathbf{x}^{\text{test}}}$ with dimension equal to the number of class labels is used to store all K results of $\mathbf{x}^{\text{test}}$, where if the class label predicted by the kth ($k \in [1, \ldots, K]$) ELM is i, the value of the corresponding entry i in the vector $\mathbf{S}_{K,\mathbf{x}^{\text{test}}}$ is increased by one. After all K results are assigned to $\mathbf{S}_{K,\mathbf{x}^{\text{test}}}$, the final class label of $\mathbf{x}^{\text{test}}$ is then determined by conducting a majority voting:

$$\varphi \left(\mathbf{x}^{\text{test}} \right) = \arg \max_{i \in [1, \ldots, C]} \mathbf{S}_{K,\mathbf{x}^{\text{test}}}, \quad (4)$$

where C is the total number of classes in the database.

The V-ELM algorithm is simple yet effective. Its algorithm structure creates a lot of rooms for further development. One such possibility is to increase the ensemble size but only select a few individual ELMs for decision making. In detail, we create an ensemble of K' ELMs and select K of them to combine the outputs. The selection is based on the mean value of norms of output weights $\|\beta\|$. Smaller $\|\beta\|$ could lead to better generalization performance [4] and this characteristic has been applied in several ELM based methods [8, 26]. The SV-ELM method is briefly described as follows.

(1) Randomly generate K' sets of hidden node parameters, train each ELM, and obtain the corresponding output weight matrix.

(2) Select K individual ELMs with small $\|\beta\|$ values in the final decision ensemble.

(3) Apply the above selected K ELM models on testing sample $\mathbf{x}^{\text{test}}$ to get the predicted label.

(4) Combine K predicted labels for the testing sample to reach the final decision.

4. Experiments

4.1. Data Collection and Processing. This was a retrospective observational study on emergency department (ED) patients. Patients were recruited at the ED of Singapore General Hospital. Eight vital signs and raw electrocardiography (ECG) data were acquired. These vital signs include temperature, respiration rate, pulse rate, systolic blood pressure (SBP), diastolic blood pressure (DBP), oxygen saturation (SpO$_2$), Glasgow Coma Scale (GCS), and pain score. ECG signals were acquired using LIFEPAK 12 defibrillator/monitor and downloaded using the CODE-STAT Suite. To ensure qualified RR intervals for calculating HRV measures, only cases containing more than 70% sinus rhythm were included in the database. Each patient was represented as a 24-dimensional feature vector (16 HRV measures and 8 vital signs) and the corresponding outcome is coded as either 0 (no cardiac arrest within 72 hours) or 1 (cardiac arrest within 72 hours). Prior to implementing ELM based risk stratification, min-max normalization [27] was performed to transform the feature set into the interval $[-1, 1]$.

4.2. Experiment Setting and Performance Evaluation. Experiments were carried out in MATLAB R2009a (Mathworks, Natick, MA) under a desktop computer with Intel 3.2 GHz CPU and 4 G RAM. The LIBSVM library [28] was used to implement linear SVM algorithm for the DIST system. The Gaussian radial basis function (RBF) activation function $G(\mathbf{w}, b, \mathbf{x}) = \exp(-\|\mathbf{x} - \mathbf{w}\|^2/b)$ was adopted for all ELM algorithms. In this study, b was chosen as 1 as a default value in MATLAB setting.

We evaluated scoring systems with the leave-one-out cross validation (LOOCV) framework. Given a dataset of K samples, one sample was selected to validate a scoring model trained with the rest of $K - 1$ samples. To complete the LOOCV based validation, all K samples had to be tested individually through K iterations. Having derived the risk scores

for all samples in the dataset, the receiver operation characteristic (ROC) analysis was conducted, with which the area under the curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were derived for performance evaluation.

4.3. Baseline Characteristics. During the recruitment period (November 2006 to December 2007), 1025 patients presented to the ED of SGH were conveniently sampled. Of these, 52 (5.1%) patients met the primary outcome (cardiac arrest within 72 hours) while 973 (94.9%) patients did not meet the primary outcome. Table 1 shows the characteristics of all recruited patients. The diagnosis group was based on physician clinical judgment. 425 (41.5%) of the patients were diagnosed under cardiovascular group, followed by 159 (15.5%) who were at respiratory group. In the primary outcome group, patients had a median age of 71 (interquartile range or IQR: 61–81), and 37 (71.2%) of patients were male.

4.4. Performance with Different Scoring Methods. Table 2 shows the comparisons of prediction results in terms of AUC, sensitivity, specificity, PPV, and NPV. The cutoff value for each scoring method was chosen as the optimal point closest to the upper left corner in the ROC curve. The 95% confidence intervals (CI) were also reported. As observed in Table 2, ELM based scoring methods outperformed SVM in risk prediction, and the proposed SV-ELM algorithm achieved the best performance. Noting that the database was fairly small and the feature dimension was not high, training time might not be a concern especially when the training process was done offline. Therefore, we briefly describe training time as follows instead of providing a detailed report. Both SVM and ELM based scoring methods completed their training in 0.3 s while V-ELM required more than 1 s and SV-ELM used more than 2 s. These data were based on the averaged value during the leave-one-out cross validation. SVM ran fast because linear kernel was implemented; no grid search [28] was needed to fine-tune parameters. As reported in [24], SVM using RBF kernel took much longer time than ELM methods for training. In the next section, we will investigate the effects of parameter setting in ELM based scoring methods.

4.5. Performance with Different ELM Parameters. In practice, the number of hidden nodes and the ensemble size usually control the network complexity and the learning performance. The V-ELM algorithm was used to illustrate the impact of parameter selection. Figures 2(a) and 2(b) depict the performance of V-ELM with different ensemble size and different number of hidden nodes, respectively. Good prediction results were obtained when the number of hidden nodes was 20 and the ensemble size was 15. This pair of parameters was observed efficient in producing a trade-off between prediction performance and system complexity. It is worth noting that the performance was more sensitive to the ensemble size compared to the number of hidden nodes.

A further investigation on the parameter K' in SV-ELM algorithm was conducted. Assuming that the base ensemble size K was 15, we gradually increased the ensemble size to 40 and only selected 15 of individual ELMs into the decision

TABLE 1: Baseline characteristics of study patients.

	No cardiac arrest within 72 h ($n = 973$)	Cardiac arrest within 72 h ($n = 52$)	P value[a]
Age (IQR)	63 (51 to 74)	71 (61 to 81)	0.007
Male gender	585 (60.1)	37 (71.2)	0.113
Race			0.080
Chinese	648 (66.6)	41 (78.8)	
Malay	147 (15.1)	6 (11.5)	
Indian	125 (12.8)	1 (1.9)	
Others	53 (5.4)	4 (7.7)	
Diagnosis grouping[b]			<0.001
Cardiovascular	413 (42.4)	12 (23.1)	
Respiratory	147 (15.1)	12 (23.1)	
Neurological	95 (9.8)	4 (7.7)	
Gastrointestinal	50 (5.1)	2 (3.8)	
Renal	28 (2.9)	0 (0)	
Endocrine	61 (6.3)	4 (7.7)	
Infectious diseases	63 (6.5)	2 (3.8)	
Vascular	22 (2.3)	3 (5.8)	
Trauma	32 (3.3)	2 (3.8)	
Cancer	30 (3.1)	10 (19.2)	
Others	32 (3.3)	1 (1.9)	
Medical history[c]			
Diabetes	323 (33.2)	15 (28.8)	0.516
Hypertension	532 (54.7)	26 (50)	0.509
Heart disease	345 (35.5)	17 (32.7)	0.684
Renal disease	127 (13.1)	11 (21.2)	0.095
Respiratory disease	116 (11.9)	7 (13.5)	0.739
Stroke	66 (6.8)	6 (11.5)	0.191
Cancer	67 (6.9)	12 (23.1)	<0.001
Others	595 (61.2)	28 (53.8)	0.293
Prior medical therapy[d]			
Beta-blockers	269 (27.6)	13 (25)	0.677
Calcium-channel blockers	189 (19.4)	11 (21.2)	0.759
Digoxin	56 (5.8)	3 (5.8)	0.997
Amiodarone	15 (1.5)	0 (0)	0.367
Other antiarrhythmics	4 (0.4)	1 (1.9)	0.127

Data shown are numbers (%) unless otherwise stated. IQR: interquartile range (25th to 75th percentiles).
[a]P value from either the chi-square test or the Mann-Whitney test as appropriate.
[b]Based on admitting emergency physician clinical diagnosis.
[c]Medical history at presentation to the emergency department.
[d]Prior outpatient medical therapy at presentation to the emergency department.

TABLE 2: Prediction results using various scoring methods. The number of hidden nodes for ELM algorithms was 20. The size of ensemble K for both V-ELM and SV-ELM algorithms was 15. The initial ensemble size K' for SV-ELM was 25.

Method	AUC (95% CI)	Cutoff	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
SVM	0.733 (0.654–0.812)	64	76.9% (65.5%–88.4%)	63.0% (60.0%–66.0%)	10.0% (7.1%–12.9%)	98.1% (97.0%–99.2%)
ELM	0.736 (0.656–0.815)	62	75.0% (63.2%–86.8%)	61.9% (58.8%–64.9%)	9.5% (6.7%–12.4%)	97.9% (96.7%–99.0%)
V-ELM	0.749 (0.671–0.827)	64	78.8% (67.7%–89.9%)	63.8% (60.8%–66.8%)	10.4% (7.4%–13.5%)	98.3% (97.3%–99.3%)
SV-ELM	0.754 (0.676–0.832)	64	78.8% (67.7%–89.9%)	64.7% (61.7%–67.8%)	10.7% (7.6%–13.8%)	98.3% (97.3%–99.3%)

(a) Number of hidden nodes = 20
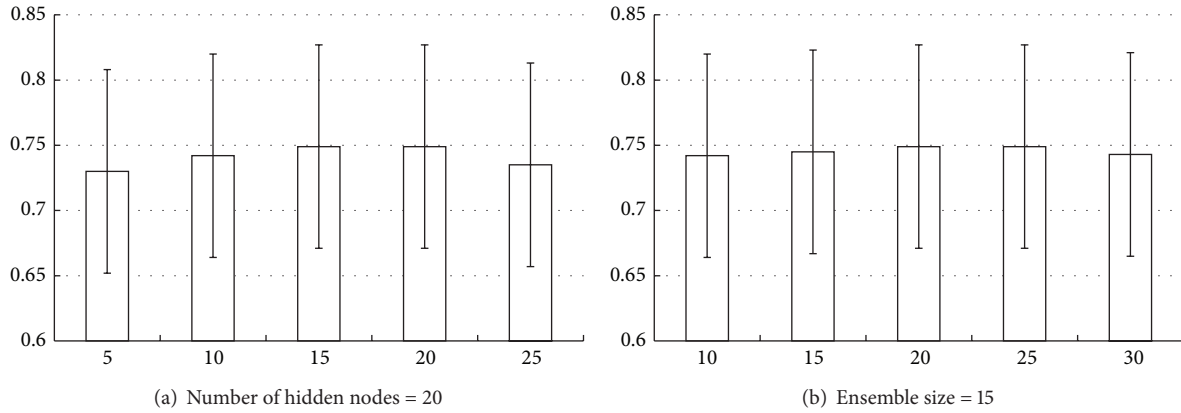
(b) Ensemble size = 15

FIGURE 2: Prediction results in terms of AUC values using V-ELM with (a) different ensemble size where the number of hidden nodes was a constant value of 20 and (b) different number of hidden nodes where the ensemble size was a constant value of 15. The 95% CIs are shown in both figures.

TABLE 3: Prediction results using SV-ELM with different initial ensemble size.

Ensemble size	AUC (95% CI)	Cutoff	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
15	0.749 (0.671–0.827)	64	78.8% (67.7%–89.9%)	63.8% (60.8%–66.8%)	10.4% (7.4%–13.5%)	98.3% (97.3%–99.3%)
15 + 5	0.753 (0.675–0.831)	64	78.8% (67.7%–89.9%)	64.5% (61.5%–67.5%)	10.6% (7.5%–13.7%)	98.3% (97.3%–99.3%)
15 + 10	0.754 (0.676–0.832)	64	78.8% (67.7%–89.9%)	64.7% (61.7%–67.8%)	10.7% (7.6%–13.8%)	98.3% (97.3%–99.3%)
15 + 15	0.746 (0.667–0.824)	64	76.9% (65.5%–88.4%)	64.9% (61.9%–67.9%)	10.5% (7.4%–13.5%)	98.1% (97.1%–99.2%)
15 + 20	0.746 (0.667–0.824)	64	76.9% (65.5%–88.4%)	64.5% (61.5%–67.5%)	10.4% (7.3%–13.4%)	98.1% (97.1%–99.2%)
15 + 25	0.735 (0.656–0.814)	64	75.0% (63.2%–86.8%)	64.7% (61.7%–67.8%)	10.2% (7.2%–13.2%)	98.0% (96.9%–99.1%)

ensemble. Table 3 presents the prediction performance. An initial ensemble size of 25 produced the best performance. With a cutoff score of 64, it achieved an AUC of 0.754, 78.8% sensitivity, 64.7% specificity, 10.7% PPV, and 98.3% NPV. We noted that a large K' dramatically reduced the prediction performance; for example, $K' = 40$ only received an AUC of 0.735. In this study, parameters K' and K were empirically selected; they were far from optimal. Therefore, derivation of a general guideline for parameter selection is worthy of further investigation.

5. Conclusions

In this retrospective observational study of 1025 critically ill patients presented to the ED, we found that ELM based methods outperformed the original SVM based risk scoring method. Furthermore, our proposed SV-ELM method achieved the best performance in ROC analysis. Based on these discoveries, we foresee the potential use of ELM methods for risk modeling in biomedical applications. ELM methods provide an alternative solution to traditional classification tools like SVM by offering an increased predictive ability.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] N. Liu, Z. Lin, J. Cao et al., "An intelligent scoring system and its application to cardiac arrest prediction," *IEEE Transactions on Information Technology in Biomedicine*, vol. 16, no. 6, pp. 1324–1331, 2012.

[2] E. M. Antman, M. Cohen, P. J. L. M. Bernink et al., "The TIMI risk score for unstable angina/non-ST elevation MI: a method for prognostication and therapeutic decision making," *Journal of the American Medical Association*, vol. 284, no. 7, pp. 835–842, 2000.

[3] C. P. Subbe, R. G. Davies, E. Williams, P. Rutherford, and L. Gemmell, "Effect of introducing the modified early warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions," *Anaesthesia*, vol. 58, no. 8, pp. 797–802, 2003.

[4] G. Huang, Q. Zhu, and C. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.

[5] G. 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] Z. Sun, K. Au, and T. Choi, "A neuro-fuzzy inference system through integration of fuzzy logic and extreme learning machines," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 37, no. 5, pp. 1321–1331, 2007.

[7] Y. Miche, A. Sorjamaa, P. Bas, O. Simula, C. Jutten, and A. Lendasse, "OP-ELM: optimally pruned extreme learning

machine," *IEEE Transactions on Neural Networks*, vol. 21, no. 1, pp. 158–162, 2010.

[8] N. Liu and H. Wang, "Ensemble based extreme learning machine," *IEEE Signal Processing Letters*, vol. 17, no. 8, pp. 754–757, 2010.

[9] J.-H. Zhai, H.-Y. Xu, and X.-Z. Wang, "Dynamic ensemble extreme learning machine based on sample entropy," *Soft Computing*, vol. 16, no. 9, pp. 1493–1502, 2012.

[10] S. Ding, X. Xu, and R. Nie, "Extreme learning machine and its applications," *Neural Computing and Applications*, vol. 25, no. 3-4, pp. 549–556, 2014.

[11] S. Suresh, R. Venkatesh Babu, and H. J. Kim, "No-reference image quality assessment using modified extreme learning machine classifier," *Applied Soft Computing Journal*, vol. 9, no. 2, pp. 541–552, 2009.

[12] N. Liu and H. Wang, "Evolutionary extreme learning machine and its application to image analysis," *Journal of Signal Processing Systems*, vol. 73, pp. 1–9, 2013.

[13] Y. Jin, J. Cao, Q. Ruan, and X. Wang, "Cross-modality 2D-3D face recognition via multiview smooth discriminant analysis based on ELM," *Journal of Electrical and Computer Engineering*, vol. 2014, Article ID 584241, 9 pages, 2014.

[14] N. Liang, P. Saratchandran, G. Huang, and N. Sundararajan, "Classification of mental tasks from EEG signals using extreme learning machine," *International Journal of Neural Systems*, vol. 16, no. 1, pp. 29–38, 2006.

[15] Y. Song, J. Crowcroft, and J. Zhang, "Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine," *Journal of Neuroscience Methods*, vol. 210, no. 2, pp. 132–146, 2012.

[16] J. Kim, H. S. Shin, K. Shin, and M. Lee, "Robust algorithm for arrhythmia classification in ECG using extreme learning machine," *Biomedical Engineering Online*, vol. 8, article 31, 2009.

[17] N. Liu, Z. Lin, Z. Koh, G. Huang, W. Ser, and M. E. H. Ong, "Patient outcome prediction with heart rate variability and vital signs," *Journal of Signal Processing Systems*, vol. 64, no. 2, pp. 265–278, 2011.

[18] R. Zhang, G.-B. Huang, N. Sundararajan, and P. Saratchandran, "Multicategory classification using an extreme learning machine for microarray gene expression cancer diagnosis," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 4, no. 3, pp. 485–494, 2007.

[19] J. Cao and L. Xiong, "Protein sequence classification with improved extreme learning machine algorithms," *BioMed Research International*, vol. 2014, Article ID 103054, 12 pages, 2014.

[20] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics B: Cybernetics*, vol. 42, no. 2, pp. 513–529, 2012.

[21] Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology, "Heart rate variability: standards of measurement, physiological interpretation, and clinical use," *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.

[22] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

[23] J. Chorowski, J. Wang, and J. M. Zurada, "Review and performance comparison of SVM- and ELM-based classifiers," *Neurocomputing*, vol. 128, pp. 507–516, 2014.

[24] J. Cao, Z. Lin, G. Huang, and N. Liu, "Voting based extreme learning machine," *Information Sciences*, vol. 185, pp. 66–77, 2012.

[25] D. Serre, *Matrices: Theory and Applications*, Springer, New York, NY, USA, 2002.

[26] Q.-Y. Zhu, A. K. Qin, P. N. Suganthan, and G.-B. Huang, "Evolutionary extreme learning machine," *Pattern Recognition*, vol. 38, no. 10, pp. 1759–1763, 2005.

[27] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2006.

[28] C. C. Chang and C. J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 1–27, 2011.

Advances in
Operations Research

Advances in
Decision Sciences

Journal of
Applied Mathematics

Algebra

Journal of
Probability and Statistics

The Scientific
World Journal

International Journal of
Differential Equations

International Journal of
Combinatorics

Advances in
Mathematical Physics

Hindawi

Submit your manuscripts at
http://www.hindawi.com

Journal of
Complex Analysis

Journal of
Mathematics

Mathematical Problems
in Engineering

Abstract and
Applied Analysis

Discrete Dynamics in
Nature and Society

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Discrete Mathematics

Journal of
Function Spaces

International Journal of
Stochastic Analysis

Journal of
Optimization