

Impact of Different Random Initializations on Generalization Performance of Extreme Learning Machine

Xiaofang Zhang¹, Xiaoli Lin¹, Rana Aamir Raza Ashfaq^{2*}

¹ Information Engineering Department, City College, Wuhan University of Science and Technology, Wuhan 430083, China.

² Department of Computer Science, Bahauddin Zakariya University, Multan, Pakistan.

* Corresponding author. Email: aamir@bzu.edu.pk

Manuscript submitted September 26, 2017; accepted November 21, 2017.

doi: 10.17706/jcp.13.7.805-822

Abstract: The generalization performance of extreme learning machine (ELM) is influenced by the random initializations to input-layer weights and hidden-layer biases. In this paper, we demonstrate this conclusion through testing the classification accuracies of ELMs corresponding to different random initializations. 30 UCI data sets and 24 continuous probability distributions are employed in this experimental study. The final results present the following important and valuable observations and conclusions, i.e., (1) the probability distributions with symmetrical and bell-shaped probability density functions (e.g., *Hyperbolic Secant*, *Student's-t*, *Laplace* and *Normal*) always bring about the higher training accuracies and easily cause the over-fitting of ELM; (2) ELMs with random input-layer weights and hidden-layer biases chosen from heavy-tailed distributions (e.g., *Gamma*, *Rayleigh* and *Frechet*) have the better generalization performances; and (3) the light-tailed distributions (e.g., *Central Chi-Squared*, *Erlang*, *F*, *Gumbel* and *Logistic*) are usually unsuited to initialize the input-layer weights and hidden-layer biases for ELM. All these provide the useful enlightenments for practical applications of ELMs in different fields.

Key words: Extreme learning machine, ELM, generalization performance, random initialization.

1. Introduction

Extreme learning machine (ELM) [1] is a kind of special single hidden-layer feed-forward neural network (SLFN) in which the input-layer weights and hidden-layer biases are randomly selected and the output-layer weights are analytically determined by solving Moore-Penrose generalized inverse [2] of hidden-layer output matrix. Because of avoiding the time-consuming iterations, ELMs obtain the extremely fast training speed. Meanwhile, the theoretical proof [3] guarantees the universal approximate capability of ELM. The lower computational complexity and acceptable generalization performance makes ELM obtain a wide range of practical applications [4].

However, ELM has an obvious defect, i.e., the instability of predictive results. Although Huang *et al.* in [1] proved the uniform convergence of ELM for any continuous probability distribution, this defect indeed exists in the practical applications due to the inherent character of data set, e.g., noise and outlier, etc. The researchers have developed some representative works to improve the instability of ELM caused by the random initializations of input-layer weights and hidden-layer biases. Zhu *et al.* [5] proposed an evolutionary ELM (E-ELM) which uses the differential evolutionary algorithm to select the input weights and hidden biases for ELM. Then, Cao *et al.* [6] improved E-ELM and developed a self-adaptive evolutionary

extreme learning machine (SaE-ELM) to optimize the hidden node parameters. Experimental results show SaE-ELM outperforms E-ELM. An optimized extreme learning machine (O-ELM) was designed by Matias *et al.* [7], which uses three different optimization algorithms to simultaneously optimize the input weights, hidden biases, and regularization factor. Heeswijk *et al.* [8] proposed two weight initialization schemes, i.e., binary ELM based on $\{0, 1\}$ -weights and ternary ELM based on $\{-1, 0, 1\}$ -weights, to improve the diversity of neurons in the hidden layer. For binary/ternary ELMs, the necessary optimizations are also required to select the better parameters. From the above-mentioned description, we can find that E-ELM, SaE-ELM, O-ELM, and binary/ternary ELMs used the different optimization algorithms to tune the initially random input-layer weights and hidden-layer biases. These works confirmed the impact of random initialization on generalization capability of ELM.

Usually, the random numbers obeying uniform distribution are used to initialize the input-layer weights and hidden-layer biases, including the original ELM [1], E-ELM [5], SaE-ELM [6], O-ELM [7], binary/ternary ELMs [8], and other forms of improved ELM [3], [9]-[15]. We rarely find ELMs which initialize the parameters with other probability distributions. The main objective of this paper is to validate the impact of different random initializations on generalization performance of ELM. Specifically, we want to clarify the influence of initializations obeying different continuous probability distributions on the training and testing accuracies of ELM. 24 different continuous probability distributions [16] as shown in Table 1 are introduced in our study to initialize the input-layer weights and hidden-layer biases for ELMs. We select 30 UCI classification data sets which involve a wide range of real domains and data characteristics. In our experiment, the 10-fold cross-validation procedure are repeated 100 times. In every time of 10-fold cross-validation, different ELMs are trained on the same training sets and evaluated on the same testing set. The training and testing accuracies on given data set are the average values of 100 times of 10-fold cross-validations. We further use Wilcoxon signed-ranks test and Friedman test [17] to compare the classification performances of different ELMs for 30 data sets. The experimental results and statistical analysis show (1) the symmetrical and bell-shaped probability distributions easily cause the over-fitting of ELM; (2) the heavy-tailed distributions bring about the better generalization performance for ELM; and (3) the light-tailed distributions obtain the worse training and testing accuracies than *Uniform* distribution. The obtained conclusions from experimental results can instruct users choose the appropriate probability distribution to initialize the necessary parameters of ELM so as to reduce the possibility of instability of ELM caused by the random initializations.

The rest of the paper is organized as follows: In Section 2, we summarize the basic extreme learning machine. In Section 3, we experimentally validate the influence of initializations obeying different probability distributions on generalization performance of ELM. We conclude this work with some remarks in the last Section.

2. Extreme Learning Machine

Here, we consider using extreme learning machine (ELM) [12] to deal with the classification problem. Given the training dataset $(X_{N \times D}, Y_{N \times C})$ and testing dataset $(X'_{M \times D}, Y'_{M \times C})$:

$$X_{N \times D} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1D} \\ x_{21} & x_{22} & \cdots & x_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{ND} \end{bmatrix}, Y_{N \times C} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1C} \\ y_{21} & y_{22} & \cdots & y_{2C} \\ \vdots & \vdots & \ddots & \vdots \\ y_{N1} & y_{N2} & \cdots & y_{NC} \end{bmatrix}$$

$$X'_{M \times D} = \begin{bmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_M \end{bmatrix} = \begin{bmatrix} x'_{11} & x'_{12} & \cdots & x'_{1D} \\ x'_{21} & x'_{22} & \cdots & x'_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ x'_{M1} & x'_{M2} & \cdots & x'_{MD} \end{bmatrix}, Y'_{M \times C} = \begin{bmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_M \end{bmatrix} = \begin{bmatrix} y'_{11} & y'_{12} & \cdots & y'_{1C} \\ y'_{21} & y'_{22} & \cdots & y'_{2C} \\ \vdots & \vdots & \ddots & \vdots \\ y'_{M1} & y'_{M2} & \cdots & y'_{MC} \end{bmatrix},$$

where N is the number of training instances, D is the number of input variables, M is the number of testing instances, and C is the number of output variables. For i -th training instance $x_i (1 \leq i \leq N)$ which belongs to c -th ($1 \leq c \leq C$) class, $y_{ic} = 1$ and $y_{ic'} = 0 (c' = 1, 2, \dots, C \text{ and } c' \neq c)$. $Y'_{M \times C}$ is unknown and needs to be predicted.

ELM is a single hidden-layer feed-forward neural network (SLFN) and does not require any iterative optimization to input/output weights. ELM determines $Y'_{M \times C}$ as follows:

Table 1. 24 Continuous Probability Distributions

Distribution	Density function $p(x)$	Support interval
D ₁ Beta	$\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, \alpha > 0, \beta > 0$	$x \in [0,1]$
D ₂ Cauchy	$\frac{1}{\pi} \left(\frac{\lambda}{x^2 + \lambda^2} \right), \lambda > 0$	$x \in (-\infty, +\infty)$
D ₃ Central Chi-Squared	$\frac{2^{-k/2} x^{k/2-1} e^{-x/2}}{\Gamma(k/2)}, k > 0$	$x \in [0, +\infty)$
D ₄ Chi	$\frac{2^{1-k/2} x^{k-1} e^{-x/2}}{\Gamma(k/2)}, k > 0$	$x \in [0, +\infty)$
D ₅ Erlang	$\frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!}, k > 0, \lambda > 0$	$x \in [0, +\infty)$
D ₆ Exponential	$\lambda \exp(-\lambda x), \lambda > 0$	$x \in [0, +\infty)$
D ₇ F	$\frac{n_1^{n_1/2} n_2^{n_2/2}}{n_1^{n_1/2} n_2^{n_2/2}} \frac{x^{n_1/2-1}}{x^{n_1/2-1}}, n_1 > 0, n_2 > 0$	$x \in [0, +\infty)$
D ₈ Frechet	$\frac{\alpha}{s} \left(\frac{x}{s} \right)^{-1-\alpha} e^{-\left(\frac{x}{s}\right)^{-\alpha}}, \alpha > 0, s > 0$	$x \in [0, +\infty)$
D ₉ Gamma	$x^{k-1} \frac{e^{-x/\theta}}{\theta^k \Gamma(k)}, k > 0, \theta > 0$	$x \in [0, +\infty)$
D ₁₀ Gumbel	$\frac{z \cdot \exp(-z)}{\beta}, z = \exp[-(x - \mu)/\beta], \mu \in R, \beta > 0$	$x \in (-\infty, +\infty)$
D ₁₁ Hyperbolic Secant	$\frac{1}{2} \operatorname{sech} \left(\frac{\pi}{2} x \right)$	$x \in (-\infty, +\infty)$
D ₁₂ Laplace	$\frac{1}{2b} \exp \left(-\frac{ x-\mu }{b} \right), \mu \in R, b > 0$	$x \in (-\infty, +\infty)$
D ₁₃ Logistic	$\frac{1}{4s} \operatorname{sech}^2 \left(\frac{x-\mu}{2s} \right), \mu \in R, s > 0$	$x \in (-\infty, +\infty)$
D ₁₄ Lognormal	$\frac{1}{x\sigma\sqrt{2\pi}} \exp \left[-\frac{(\ln x - \mu)^2}{2\sigma^2} \right], \mu \in R, \sigma^2 > 0$	$x \in (0, +\infty)$
D ₁₅ Normal	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(x-\mu)^2}{2\sigma^2} \right], \mu \in R, \sigma^2 > 0$	$x \in (0, +\infty)$
D ₁₆ Pareto	$\frac{\alpha x_m^\alpha}{x^{\alpha+1}}, x_m > 0, \alpha > 0$	$x \in [x_m, +\infty)$
D ₁₇ Rayleigh	$\frac{x}{\sigma^2} \exp \left(-\frac{x^2}{2\sigma^2} \right), \sigma > 0$	$x \in [0, +\infty)$
D ₁₈ Semicircle	$\frac{2}{\pi r^2} \sqrt{r^2 - x^2}, r > 0$	$x \in [-r, r]$
D ₁₉ Student's-t	$\frac{(1+x^2/v)^{-v/2}}{\sqrt{v} B(1/2, v/2)}, v > 0$	$x \in (-\infty, +\infty)$
D ₂₀ Triangular	$\begin{cases} \frac{2(x-a)}{(b-a)(c-a)}, & a \leq x \leq c \\ \frac{2(b-x)}{(b-a)(b-c)}, & c \leq x \leq b \end{cases}$	$x \in [a, b]$
D ₂₁ Truncated Normal	$\frac{\frac{1}{\sigma} \phi \left(\frac{x-\mu}{\sigma} \right)}{\Phi \left(\frac{b-\mu}{\sigma} \right) - \Phi \left(\frac{a-\mu}{\sigma} \right)}, \mu \in R, a \in R, b \in R, \sigma^2 > 0$	$x \in [a, b]$
D ₂₂ Uniform	$\frac{1}{b-a}$	$x \in [a, b]$
D ₂₃ Von Mises	$\frac{e^{k \cos(x-\mu)}}{2\pi I_0(k)}, \mu \in R, k > 0$	$x \in [-\pi, \pi]$
D ₂₄ Weibull	$\frac{k}{\lambda} \left(\frac{x}{\lambda} \right)^{k-1} e^{-(x/\lambda)^k}, \lambda > 0, k > 0$	$x \in [0, +\infty)$

Note: $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$ is the Beta function. $\Psi(x) = d\ln[\Gamma(x)]/dx$ is the Digamma function, where $\Gamma(x) = \int_0^\infty u^{x-1}e^{-u} du$ is the Gamma function. $I_0(x)$ is the modified Bessel function of order 0, $I_1(x)$ and is the modified Bessel function of order 1.

$$Y'_{M \times C} = H'_{M \times L} \beta_{L \times C} = H'_{M \times L} H_{L \times N}^\dagger Y_{N \times C} = \begin{cases} H'(H^T H)^{-1} H^T Y, & \text{if } N \geq L \\ H' H^T (H^T)^{-1} Y, & \text{if } N < L \end{cases} \quad (1)$$

Table 2. The Details of 30 UCI Data Sets

	Data sets	Attributes	Classes	Class distribution	Instances
1	Auto Mpg	5	3	245/79/68	392
2	Blood Transfusion	4	2	570/178	748
3	Breast Cancer	10	2	458/241	699
4	Breast Cancer W-D	30	2	357/212	569
5	Breast Cancer W-P	33	2	151/47	198
6	Cleveland	13	5	160/54/35/35/13	297
7	Credit Approval	15	2	383/307	690
8	Cylinder Bands	20	2	312/228	540
9	Ecoli	5	8	143/77/52/35/20/5/2/2	336
10	Glass Identification	9	7	76/70/29/17/13/9/0	214
11	Haberman's Survival	3	2	225/81	306
12	Heart Disease	13	2	150/120	270
13	Image Segment	19	7	330 × 7	2310
14	Ionosphere	33	2	225/126	351
15	Iris	4	3	50 × 3	150
16	Libras Movement	90	15	24 × 15	360
17	Magic Telescope	10	2	12332/6688	19020(10%)
18	Musk Version 1	166	2	269/207	476
19	New Thyroid Gland	5	3	150/35/30	215
20	Page Blocks	10	5	4913/329/115/88/28	5473(10%)
21	Parkinsons	22	2	147/48	195
22	Pima Indian Diabetes	8	2	500/268	768
23	Sonar	60	2	111/97	208
24	SPECTF Heart	44	2	212/55	267
25	Vehicle Silhouettes	18	4	218/217/212/199	846
26	Vowel Recognition	10	11	48 × 11	528
27	Wine	13	3	91/59/48	178
28	Wine Quality-Red	11	6	681/638/199/53/18/10	1599
29	Wine Quality-White	11	6	2198/1457/880/175/163/20	4898(10%)
30	Yeast	8	10	463/429/244/163/51/44/35/30/20/5	1484

where $\beta_{L \times C}$ is the output-layer weights, L is the number of hidden-layer nodes

$$H_{N \times L} = \begin{bmatrix} h(x_1) \\ h(x_2) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} g(w_1 x_1 + b_1) & g(w_2 x_1 + b_2) & \cdots & g(w_L x_1 + b_L) \\ g(w_1 x_2 + b_1) & g(w_2 x_2 + b_2) & \cdots & g(w_L x_2 + b_L) \\ \vdots & \vdots & \ddots & \vdots \\ g(w_1 x_N + b_1) & g(w_2 x_N + b_2) & \cdots & g(w_L x_N + b_L) \end{bmatrix} \quad (2)$$

is the hidden-layer output matrix for training instances, $H_{L \times N}^\dagger$ is Moore-Penrose generalized inverse of $H_{N \times L}$,

$$H_{M \times L}' = \begin{bmatrix} h(x_1') \\ h(x_2') \\ \vdots \\ h(x_M') \end{bmatrix} = \begin{bmatrix} g(w_1 x_1' + b_1) & g(w_2 x_1' + b_2) & \cdots & g(w_L x_1' + b_L) \\ g(w_1 x_2' + b_1) & g(w_2 x_2' + b_2) & \cdots & g(w_L x_2' + b_L) \\ \vdots & \vdots & \ddots & \vdots \\ g(w_1 x_M' + b_1) & g(w_2 x_M' + b_2) & \cdots & g(w_L x_M' + b_L) \end{bmatrix} \quad (3)$$

Table 3. Training Accuracies of ELMs Corresponding to 24 Probability Distributions on 30 UCI Data Sets

	D ₁ :Beta	D ₂ :Cauchy	D ₃ :Central Chi	D ₄ :Chi	D ₅ :Erlang	D ₆ :Exponential
1	0.858 ± 0.002 (1.5)	0.848 ± 0.003 (1.5)	0.863 ± 0.003 (7.5)	0.867 ± 0.002 (24.0)	0.867 ± 0.003 (5.0)	0.862 ± 0.002 (5.0)
2	0.806 ± 0.001 (15.0)	0.813 ± 0.001 (7.5)	0.812 ± 0.001 (13.0)	0.810 ± 0.001 (2.0)	0.812 ± 0.001 (21.5)	0.811 ± 0.001 (21.5)
3	0.975 ± 0.001 (5.0)	0.975 ± 0.001 (13.0)	0.977 ± 0.001 (5.0)	0.978 ± 0.001 (16.0)	0.975 ± 0.001 (2.0)	0.978 ± 0.001 (2.0)
4	0.976 ± 0.002 (17.0)	0.954 ± 0.004 (23.0)	0.849 ± 0.004 (14.5)	0.905 ± 0.003 (24.0)	0.842 ± 0.003 (13.0)	0.935 ± 0.003 (14.5)
5	0.840 ± 0.006 (17.0)	0.847 ± 0.007 (23.0)	0.777 ± 0.001 (13.5)	0.799 ± 0.002 (24.0)	0.773 ± 0.002 (13.5)	0.818 ± 0.003 (15.0)
6	0.697 ± 0.004 (15.0)	0.683 ± 0.004 (19.5)	0.680 ± 0.004 (12.5)	0.693 ± 0.003 (24.0)	0.681 ± 0.003 (11.0)	0.698 ± 0.004 (12.5)
7	0.804 ± 0.002 (10.0)	0.809 ± 0.003 (2.5)	0.812 ± 0.002 (9.0)	0.807 ± 0.002 (2.5)	0.812 ± 0.002 (13.0)	0.808 ± 0.002 (13.0)
8	0.708 ± 0.005 (17.0)	0.703 ± 0.006 (24.0)	0.599 ± 0.003 (15.5)	0.647 ± 0.005 (23.0)	0.589 ± 0.001 (14.0)	0.682 ± 0.004 (13.0)
9	0.904 ± 0.002 (17.0)	0.899 ± 0.002 (20.5)	0.895 ± 0.002 (12.5)	0.897 ± 0.002 (24.0)	0.894 ± 0.002 (12.5)	0.900 ± 0.001 (14.5)
10	0.816 ± 0.005 (10.0)	0.813 ± 0.006 (18.5)	0.811 ± 0.006 (10.0)	0.818 ± 0.004 (24.0)	0.815 ± 0.004 (10.0)	0.818 ± 0.005 (10.0)
11	0.784 ± 0.002 (18.0)	0.800 ± 0.003 (10.5)	0.798 ± 0.003 (16.0)	0.791 ± 0.002 (17.0)	0.799 ± 0.003 (20.5)	0.795 ± 0.003 (22.5)
12	0.885 ± 0.004 (18.0)	0.884 ± 0.005 (22.5)	0.812 ± 0.004 (13.0)	0.830 ± 0.004 (24.0)	0.800 ± 0.005 (15.0)	0.864 ± 0.005 (14.0)
13	0.929 ± 0.001 (16.0)	0.909 ± 0.003 (22.0)	0.880 ± 0.004 (15.0)	0.910 ± 0.002 (24.0)	0.875 ± 0.004 (5.0)	0.919 ± 0.002 (5.0)
14	0.861 ± 0.003 (19.0)	0.897 ± 0.007 (23.5)	0.643 ± 0.001 (13.0)	0.652 ± 0.002 (22.0)	0.641 ± 0.000 (14.0)	0.739 ± 0.007 (15.0)
15	0.988 ± 0.001 (20.0)	0.984 ± 0.002 (5.0)	0.986 ± 0.001 (20.0)	0.985 ± 0.001 (5.0)	0.987 ± 0.001 (20.0)	0.985 ± 0.001 (20.0)
16	0.862 ± 0.003 (19.0)	0.894 ± 0.007 (23.5)	0.643 ± 0.001 (13.0)	0.652 ± 0.003 (22.0)	0.641 ± 0.000 (14.0)	0.736 ± 0.007 (15.0)
17	0.851 ± 0.001 (7.5)	0.844 ± 0.002 (21.0)	0.843 ± 0.002 (6.0)	0.848 ± 0.002 (24.0)	0.840 ± 0.002 (3.0)	0.850 ± 0.001 (3.0)
18	0.565 ± 0.000 (17.5)	0.752 ± 0.008 (17.5)	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (17.5)
19	0.952 ± 0.001 (12.0)	0.959 ± 0.001 (4.5)	0.955 ± 0.001 (12.0)	0.954 ± 0.001 (1.0)	0.955 ± 0.001 (20.5)	0.954 ± 0.001 (20.5)
20	0.970 ± 0.001 (3.0)	0.973 ± 0.002 (14.5)	0.972 ± 0.001 (1.0)	0.974 ± 0.001 (20.5)	0.971 ± 0.001 (3.0)	0.975 ± 0.001 (6.5)
21	0.979 ± 0.003 (6.5)	0.937 ± 0.005 (21.0)	0.941 ± 0.005 (6.5)	0.975 ± 0.003 (24.0)	0.936 ± 0.005 (3.0)	0.975 ± 0.003 (3.0)
22	0.803 ± 0.001 (17.0)	0.796 ± 0.002 (21.5)	0.787 ± 0.003 (15.0)	0.795 ± 0.001 (24.0)	0.784 ± 0.002 (13.0)	0.798 ± 0.002 (13.0)
23	0.824 ± 0.006 (17.0)	0.821 ± 0.008 (22.5)	0.534 ± 0.000 (13.0)	0.541 ± 0.003 (22.5)	0.534 ± 0.000 (14.0)	0.642 ± 0.007 (15.0)
24	0.794 ± 0.000 (17.5)	0.796 ± 0.002 (17.5)	0.794 ± 0.000 (17.5)	0.794 ± 0.000 (17.5)	0.794 ± 0.000 (17.5)	0.794 ± 0.000 (17.5)
25	0.791 ± 0.003 (16.0)	0.746 ± 0.006 (21.0)	0.726 ± 0.006 (12.5)	0.762 ± 0.004 (24.0)	0.723 ± 0.004 (4.0)	0.771 ± 0.004 (4.0)
26	0.866 ± 0.004 (16.0)	0.806 ± 0.008 (21.0)	0.761 ± 0.008 (15.0)	0.843 ± 0.006 (24.0)	0.754 ± 0.009 (14.0)	0.846 ± 0.005 (12.5)
27	0.999 ± 0.001	0.995 ± 0.002	0.939 ± 0.004	0.973 ± 0.002	0.926 ± 0.005	0.989 ± 0.002

	(17.0)	(21.0)	(13.0)	(24.0)	(14.5)	(14.5)
28	0.628 ± 0.002 (9.0)	0.623 ± 0.003 (20.0)	0.619 ± 0.002 (9.0)	0.628 ± 0.002 (24.0)	0.621 ± 0.002 (2.5)	0.628 ± 0.002 (2.5)
29	0.635 ± 0.003 (3.5)	0.634 ± 0.005 (7.5)	0.639 ± 0.005 (3.5)	0.643 ± 0.004 (24.0)	0.640 ± 0.004 (7.5)	0.643 ± 0.004 (5.0)
30	0.625 ± 0.001 (17.5)	0.623 ± 0.002 (21.0)	0.616 ± 0.002 (12.5)	0.620 ± 0.001 (24.0)	0.614 ± 0.002 (15.0)	0.623 ± 0.002 (15.0)
Rank	11.183	13.700	16.467	13.717	17.050	11.900
	D ₇ :F	D ₈ :Frechet	D ₉ :Gamma	D ₁₀ :Gumbel	D ₁₁ :Hyperbolic	D ₁₂ :Laplace
1	0.847 ± 0.003 (9.5)	0.863 ± 0.002 (21.5)	0.863 ± 0.002 (21.5)	0.861 ± 0.003 (3.0)	0.853 ± 0.003 (7.5)	0.853 ± 0.002 (19.0)
2	0.813 ± 0.001 (2.0)	0.806 ± 0.001 (13.0)	0.806 ± 0.001 (7.5)	0.813 ± 0.001 (7.5)	0.811 ± 0.001 (7.5)	0.812 ± 0.001 (13.0)
3	0.974 ± 0.001 (10.5)	0.979 ± 0.001 (20.5)	0.979 ± 0.001 (16.0)	0.976 ± 0.001 (10.5)	0.973 ± 0.001 (5.0)	0.974 ± 0.001 (20.5)
4	0.720 ± 0.039 (21.0)	0.936 ± 0.002 (3.0)	0.935 ± 0.002 (3.0)	0.848 ± 0.005 (22.0)	0.977 ± 0.001 (18.0)	0.977 ± 0.001 (3.0)
5	0.764 ± 0.001 (19.5)	0.818 ± 0.003 (4.0)	0.817 ± 0.003 (8.0)	0.779 ± 0.002 (21.0)	0.888 ± 0.005 (18.0)	0.883 ± 0.005 (5.5)
6	0.653 ± 0.004 (23.0)	0.699 ± 0.004 (5.5)	0.698 ± 0.004 (9.5)	0.675 ± 0.004 (22.0)	0.703 ± 0.004 (17.0)	0.701 ± 0.004 (5.5)
7	0.812 ± 0.003 (6.0)	0.805 ± 0.002 (17.0)	0.805 ± 0.002 (13.0)	0.811 ± 0.002 (6.0)	0.804 ± 0.002 (2.5)	0.805 ± 0.002 (19.5)
8	0.590 ± 0.003 (19.0)	0.688 ± 0.005 (2.0)	0.689 ± 0.004 (6.5)	0.601 ± 0.003 (22.0)	0.722 ± 0.005 (18.0)	0.719 ± 0.006 (4.5)
9	0.886 ± 0.003 (22.5)	0.900 ± 0.002 (5.0)	0.899 ± 0.001 (9.0)	0.893 ± 0.002 (22.5)	0.905 ± 0.002 (17.0)	0.904 ± 0.002 (5.0)
10	0.785 ± 0.005 (23.0)	0.818 ± 0.005 (2.5)	0.818 ± 0.004 (2.5)	0.804 ± 0.004 (20.0)	0.820 ± 0.006 (14.5)	0.820 ± 0.006 (5.5)
11	0.793 ± 0.003 (7.5)	0.784 ± 0.002 (6.0)	0.782 ± 0.002 (3.0)	0.801 ± 0.003 (7.5)	0.803 ± 0.002 (14.5)	0.806 ± 0.003 (5.0)
12	0.795 ± 0.006 (21.0)	0.856 ± 0.003 (2.0)	0.857 ± 0.004 (4.0)	0.809 ± 0.006 (22.5)	0.889 ± 0.004 (17.0)	0.888 ± 0.004 (6.5)
13	0.805 ± 0.010 (23.0)	0.922 ± 0.001 (9.5)	0.922 ± 0.001 (9.5)	0.866 ± 0.005 (21.0)	0.921 ± 0.002 (18.5)	0.921 ± 0.002 (5.0)
14	0.642 ± 0.001 (18.0)	0.701 ± 0.004 (1.0)	0.693 ± 0.004 (8.0)	0.653 ± 0.003 (20.0)	0.936 ± 0.003 (17.0)	0.931 ± 0.004 (4.5)
15	0.987 ± 0.002 (20.0)	0.985 ± 0.001 (12.0)	0.985 ± 0.001 (12.0)	0.985 ± 0.001 (20.0)	0.986 ± 0.001 (20.0)	0.986 ± 0.001 (12.0)
16	0.642 ± 0.001 (18.0)	0.703 ± 0.005 (1.0)	0.693 ± 0.004 (8.5)	0.654 ± 0.003 (20.0)	0.936 ± 0.003 (17.0)	0.930 ± 0.004 (3.5)
17	0.830 ± 0.004 (22.0)	0.851 ± 0.001 (11.5)	0.851 ± 0.001 (11.5)	0.838 ± 0.002 (23.0)	0.847 ± 0.001 (7.5)	0.847 ± 0.001 (16.5)
18	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (4.0)	0.565 ± 0.000 (9.0)	0.565 ± 0.000 (17.5)	0.809 ± 0.007 (17.5)	0.798 ± 0.008 (6.0)
19	0.960 ± 0.001 (12.0)	0.953 ± 0.001 (12.0)	0.953 ± 0.001 (12.0)	0.954 ± 0.001 (12.0)	0.954 ± 0.001 (3.0)	0.954 ± 0.001 (12.0)
20	0.970 ± 0.001 (10.5)	0.974 ± 0.001 (20.5)	0.973 ± 0.001 (14.5)	0.972 ± 0.001 (10.5)	0.970 ± 0.001 (3.0)	0.971 ± 0.001 (20.5)
21	0.879 ± 0.007 (23.0)	0.981 ± 0.003 (12.0)	0.981 ± 0.003 (10.0)	0.929 ± 0.005 (22.0)	0.946 ± 0.004 (9.0)	0.947 ± 0.006 (14.5)
22	0.777 ± 0.003 (23.0)	0.799 ± 0.001 (7.5)	0.799 ± 0.001 (11.0)	0.783 ± 0.003 (21.5)	0.802 ± 0.002 (18.0)	0.801 ± 0.002 (7.5)
23	0.534 ± 0.000 (19.0)	0.593 ± 0.008 (6.0)	0.573 ± 0.005 (8.0)	0.537 ± 0.003 (20.0)	0.891 ± 0.008 (18.0)	0.881 ± 0.007 (4.5)
24	0.794 ± 0.000 (17.5)	0.794 ± 0.000 (5.5)	0.794 ± 0.000 (8.0)	0.794 ± 0.000 (17.5)	0.824 ± 0.005 (17.5)	0.818 ± 0.004 (4.0)
25	0.669 ± 0.008 (23.0)	0.781 ± 0.003 (6.0)	0.781 ± 0.003 (9.5)	0.710 ± 0.005 (22.0)	0.774 ± 0.004 (18.0)	0.772 ± 0.004 (7.5)
26	0.599 ± 0.015 (23.0)	0.857 ± 0.004 (1.5)	0.858 ± 0.004 (3.5)	0.710 ± 0.009 (22.0)	0.868 ± 0.005 (18.0)	0.867 ± 0.004 (5.5)
27	0.874 ± 0.012 (22.0)	0.988 ± 0.001 (5.0)	0.988 ± 0.001 (5.0)	0.921 ± 0.007 (23.0)	1.000 ± 0.000 (18.0)	1.000 ± 0.000 (5.0)
28	0.601 ± 0.003	0.629 ± 0.002	0.629 ± 0.002	0.615 ± 0.003	0.628 ± 0.002	0.628 ± 0.002

	(23.0)	(9.0)	(9.0)	(22.0)	(18.5)	(9.0)
29	0.604 ± 0.006 (23.0)	0.640 ± 0.005 (17.5)	0.642 ± 0.004 (11.5)	0.631 ± 0.005 (20.0)	0.636 ± 0.005 (2.0)	0.638 ± 0.005 (15.0)
30	0.601 ± 0.003 (22.5)	0.622 ± 0.001 (4.5)	0.622 ± 0.001 (1.0)	0.613 ± 0.002 (22.5)	0.625 ± 0.002 (17.5)	0.626 ± 0.001 (4.5)
Rank	19.883	11.767	12.067	18.150	8.267	8.833
	D ₁₃ :Logistic	D ₁₄ :Lognormal	D ₁₅ :Normal	D ₁₆ :Pareto	D ₁₇ :Rayleigh	D ₁₈ :Semicircle
1	0.865 ± 0.003 (13.0)	0.862 ± 0.003 (23.0)	0.854 ± 0.003 (5.0)	0.859 ± 0.003 (11.5)	0.861 ± 0.003 (9.5)	0.855 ± 0.002 (16.0)
2	0.812 ± 0.001 (21.5)	0.812 ± 0.001 (2.0)	0.811 ± 0.001 (7.5)	0.807 ± 0.001 (17.5)	0.807 ± 0.001 (17.5)	0.806 ± 0.001 (21.5)
3	0.976 ± 0.001 (13.0)	0.978 ± 0.001 (13.0)	0.973 ± 0.001 (8.0)	0.977 ± 0.001 (8.0)	0.977 ± 0.001 (8.0)	0.973 ± 0.001 (20.5)
4	0.846 ± 0.004 (7.0)	0.885 ± 0.005 (12.0)	0.977 ± 0.001 (20.0)	0.882 ± 0.003 (19.0)	0.965 ± 0.002 (11.0)	0.975 ± 0.001 (9.5)
5	0.778 ± 0.002 (11.0)	0.792 ± 0.002 (10.0)	0.886 ± 0.006 (22.0)	0.779 ± 0.001 (19.5)	0.825 ± 0.005 (12.0)	0.892 ± 0.004 (1.5)
6	0.679 ± 0.004 (14.0)	0.685 ± 0.004 (18.0)	0.703 ± 0.003 (21.0)	0.681 ± 0.004 (19.5)	0.702 ± 0.003 (7.5)	0.706 ± 0.004 (1.5)
7	0.811 ± 0.002 (17.0)	0.812 ± 0.002 (8.0)	0.803 ± 0.003 (2.5)	0.811 ± 0.002 (6.0)	0.805 ± 0.002 (13.0)	0.802 ± 0.002 (22.0)
8	0.597 ± 0.002 (10.0)	0.629 ± 0.005 (11.0)	0.721 ± 0.005 (20.5)	0.599 ± 0.002 (20.5)	0.702 ± 0.005 (12.0)	0.721 ± 0.006 (4.5)
9	0.893 ± 0.002 (9.0)	0.897 ± 0.002 (14.5)	0.905 ± 0.002 (19.0)	0.894 ± 0.001 (20.5)	0.904 ± 0.002 (9.0)	0.906 ± 0.001 (2.0)
10	0.814 ± 0.005 (16.5)	0.817 ± 0.005 (21.0)	0.819 ± 0.006 (22.0)	0.820 ± 0.004 (2.5)	0.817 ± 0.005 (14.5)	0.816 ± 0.005 (16.5)
11	0.801 ± 0.003 (20.5)	0.796 ± 0.002 (9.0)	0.805 ± 0.003 (12.5)	0.781 ± 0.003 (24.0)	0.787 ± 0.002 (19.0)	0.798 ± 0.003 (12.5)
12	0.800 ± 0.006 (9.5)	0.832 ± 0.005 (11.0)	0.887 ± 0.004 (20.0)	0.817 ± 0.004 (19.0)	0.879 ± 0.004 (12.0)	0.887 ± 0.004 (6.5)
13	0.876 ± 0.005 (1.0)	0.901 ± 0.003 (17.0)	0.922 ± 0.002 (20.0)	0.901 ± 0.002 (18.5)	0.927 ± 0.001 (2.0)	0.920 ± 0.001 (13.0)
14	0.648 ± 0.001 (11.0)	0.659 ± 0.003 (10.0)	0.934 ± 0.004 (21.0)	0.641 ± 0.000 (23.5)	0.846 ± 0.003 (12.0)	0.934 ± 0.003 (4.5)
15	0.985 ± 0.002 (1.5)	0.985 ± 0.002 (24.0)	0.986 ± 0.002 (12.0)	0.988 ± 0.001 (1.5)	0.986 ± 0.001 (12.0)	0.987 ± 0.001 (5.0)
16	0.648 ± 0.002 (11.0)	0.659 ± 0.003 (10.0)	0.935 ± 0.004 (21.0)	0.641 ± 0.000 (23.5)	0.846 ± 0.003 (12.0)	0.933 ± 0.003 (6.0)
17	0.836 ± 0.002 (3.0)	0.848 ± 0.002 (19.0)	0.846 ± 0.001 (20.0)	0.847 ± 0.002 (11.5)	0.851 ± 0.001 (3.0)	0.846 ± 0.001 (16.5)
18	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (10.0)	0.806 ± 0.007 (17.5)	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (17.5)	0.823 ± 0.007 (2.0)
19	0.954 ± 0.001 (23.0)	0.956 ± 0.001 (2.0)	0.954 ± 0.001 (4.5)	0.954 ± 0.001 (12.0)	0.953 ± 0.001 (20.5)	0.954 ± 0.001 (12.0)
20	0.972 ± 0.001 (20.5)	0.974 ± 0.001 (6.5)	0.970 ± 0.001 (10.5)	0.973 ± 0.001 (6.5)	0.972 ± 0.001 (10.5)	0.970 ± 0.001 (20.5)
21	0.933 ± 0.005 (5.0)	0.962 ± 0.004 (20.0)	0.945 ± 0.005 (17.5)	0.973 ± 0.003 (8.0)	0.981 ± 0.002 (3.0)	0.941 ± 0.003 (17.5)
22	0.784 ± 0.002 (2.5)	0.793 ± 0.002 (16.0)	0.802 ± 0.002 (20.0)	0.792 ± 0.002 (19.0)	0.802 ± 0.002 (7.5)	0.803 ± 0.002 (2.5)
23	0.535 ± 0.001 (10.0)	0.540 ± 0.003 (11.0)	0.892 ± 0.007 (22.5)	0.534 ± 0.000 (22.5)	0.765 ± 0.005 (12.0)	0.895 ± 0.007 (3.0)
24	0.794 ± 0.000 (17.5)	0.794 ± 0.000 (10.0)	0.825 ± 0.004 (17.5)	0.794 ± 0.000 (17.5)	0.794 ± 0.000 (17.5)	0.842 ± 0.005 (1.0)
25	0.712 ± 0.006 (1.0)	0.748 ± 0.004 (19.0)	0.773 ± 0.004 (20.0)	0.754 ± 0.004 (17.0)	0.786 ± 0.004 (2.0)	0.771 ± 0.004 (12.5)
26	0.739 ± 0.009 (5.5)	0.811 ± 0.007 (19.0)	0.866 ± 0.004 (20.0)	0.836 ± 0.006 (17.0)	0.862 ± 0.005 (11.0)	0.865 ± 0.005 (7.5)
27	0.920 ± 0.006 (10.0)	0.967 ± 0.003 (12.0)	1.000 ± 0.000 (20.0)	0.966 ± 0.002 (19.0)	0.998 ± 0.001 (11.0)	1.000 ± 0.000 (5.0)
28	0.617 ± 0.002 (9.0)	0.623 ± 0.002 (18.5)	0.628 ± 0.002 (21.0)	0.626 ± 0.002 (16.5)	0.628 ± 0.002 (9.0)	0.626 ± 0.002 (16.5)
29	0.635 ± 0.005	0.645 ± 0.004	0.637 ± 0.004	0.646 ± 0.004	0.638 ± 0.004	0.636 ± 0.004

	(20.0)	(22.0)	(9.0)	(1.0)	(11.5)	(17.5)
30	0.613 ± 0.002 (4.5)	0.620 ± 0.001 (12.5)	0.625 ± 0.001 (20.0)	0.619 ± 0.001 (19.0)	0.624 ± 0.001 (9.5)	0.624 ± 0.002 (9.5)
Rank	18.100	13.817	8.983	15.283	10.950	10.217
	D ₁₉ :Studenti-s-t	D ₂₀ :Triangular	D ₂₁ :Truncated	D ₂₂ :Uniform	D ₂₃ :Von Mises	D ₂₄ :Weibull
1	0.854 ± 0.002 (19.0)	0.855 ± 0.003 (16.0)	0.856 ± 0.002 (14.0)	0.855 ± 0.002 (16.0)	0.854 ± 0.002 (19.0)	0.859 ± 0.002 (11.5)
2	0.812 ± 0.001 (7.5)	0.812 ± 0.001 (7.5)	0.807 ± 0.001 (17.5)	0.812 ± 0.001 (7.5)	0.807 ± 0.001 (17.5)	0.805 ± 0.001 (24.0)
3	0.973 ± 0.001 (20.5)	0.973 ± 0.001 (20.5)	0.973 ± 0.001 (20.5)	0.974 ± 0.001 (16.0)	0.972 ± 0.001 (24.0)	0.979 ± 0.001 (2.0)
4	0.977 ± 0.001 (3.0)	0.977 ± 0.001 (3.0)	0.975 ± 0.001 (9.5)	0.976 ± 0.001 (7.0)	0.976 ± 0.001 (7.0)	0.933 ± 0.002 (16.0)
5	0.885 ± 0.005 (7.0)	0.886 ± 0.006 (5.5)	0.892 ± 0.005 (1.5)	0.881 ± 0.006 (9.0)	0.891 ± 0.004 (3.0)	0.811 ± 0.002 (16.0)
6	0.702 ± 0.005 (7.5)	0.704 ± 0.004 (4.0)	0.705 ± 0.003 (3.0)	0.701 ± 0.004 (9.5)	0.706 ± 0.004 (1.5)	0.688 ± 0.004 (16.0)
7	0.805 ± 0.002 (13.0)	0.802 ± 0.002 (22.0)	0.802 ± 0.002 (22.0)	0.800 ± 0.002 (24.0)	0.803 ± 0.002 (19.5)	0.804 ± 0.002 (17.0)
8	0.719 ± 0.005 (6.5)	0.718 ± 0.005 (8.0)	0.722 ± 0.005 (2.0)	0.714 ± 0.006 (9.0)	0.722 ± 0.006 (2.0)	0.682 ± 0.004 (15.5)
9	0.904 ± 0.002 (9.0)	0.905 ± 0.002 (5.0)	0.906 ± 0.001 (2.0)	0.904 ± 0.001 (9.0)	0.906 ± 0.001 (2.0)	0.897 ± 0.002 (17.0)
10	0.820 ± 0.005 (2.5)	0.819 ± 0.005 (5.5)	0.815 ± 0.005 (18.5)	0.818 ± 0.005 (10.0)	0.818 ± 0.005 (10.0)	0.818 ± 0.005 (10.0)
11	0.806 ± 0.002 (3.0)	0.806 ± 0.003 (3.0)	0.799 ± 0.002 (10.5)	0.807 ± 0.002 (1.0)	0.796 ± 0.003 (14.5)	0.782 ± 0.002 (22.5)
12	0.889 ± 0.004 (2.0)	0.887 ± 0.004 (6.5)	0.887 ± 0.003 (6.5)	0.885 ± 0.004 (9.5)	0.889 ± 0.004 (2.0)	0.853 ± 0.004 (16.0)
13	0.922 ± 0.002 (5.0)	0.922 ± 0.001 (5.0)	0.920 ± 0.001 (13.0)	0.921 ± 0.002 (9.5)	0.920 ± 0.002 (13.0)	0.921 ± 0.001 (9.5)
14	0.932 ± 0.004 (7.0)	0.934 ± 0.004 (4.5)	0.934 ± 0.003 (4.5)	0.930 ± 0.004 (9.0)	0.935 ± 0.003 (2.0)	0.679 ± 0.002 (16.0)
15	0.986 ± 0.001 (12.0)	0.986 ± 0.001 (12.0)	0.987 ± 0.001 (5.0)	0.986 ± 0.001 (12.0)	0.987 ± 0.001 (5.0)	0.986 ± 0.001 (12.0)
16	0.932 ± 0.004 (7.0)	0.935 ± 0.004 (3.5)	0.935 ± 0.003 (3.5)	0.930 ± 0.005 (8.5)	0.935 ± 0.003 (3.5)	0.678 ± 0.002 (16.0)
17	0.847 ± 0.001 (11.5)	0.847 ± 0.001 (11.5)	0.846 ± 0.001 (16.5)	0.846 ± 0.001 (16.5)	0.847 ± 0.001 (11.5)	0.851 ± 0.001 (3.0)
18	0.799 ± 0.009 (7.5)	0.807 ± 0.009 (5.0)	0.823 ± 0.006 (2.0)	0.799 ± 0.006 (7.5)	0.823 ± 0.007 (2.0)	0.565 ± 0.000 (17.5)
19	0.954 ± 0.001 (12.0)	0.954 ± 0.001 (12.0)	0.954 ± 0.001 (12.0)	0.953 ± 0.001 (20.5)	0.954 ± 0.001 (12.0)	0.951 ± 0.001 (24.0)
20	0.971 ± 0.001 (14.5)	0.970 ± 0.001 (20.5)	0.970 ± 0.001 (20.5)	0.971 ± 0.001 (14.5)	0.970 ± 0.001 (20.5)	0.973 ± 0.001 (6.5)
21	0.946 ± 0.004 (12.0)	0.945 ± 0.006 (14.5)	0.941 ± 0.004 (17.5)	0.946 ± 0.004 (12.0)	0.941 ± 0.004 (17.5)	0.982 ± 0.003 (1.0)
22	0.802 ± 0.002 (7.5)	0.802 ± 0.002 (7.5)	0.803 ± 0.002 (2.5)	0.802 ± 0.002 (7.5)	0.803 ± 0.002 (2.5)	0.799 ± 0.002 (13.0)
23	0.888 ± 0.007 (7.0)	0.892 ± 0.007 (4.5)	0.897 ± 0.006 (1.5)	0.880 ± 0.006 (9.0)	0.897 ± 0.007 (1.5)	0.563 ± 0.001 (16.0)
24	0.820 ± 0.005 (7.0)	0.824 ± 0.004 (5.5)	0.840 ± 0.005 (2.5)	0.816 ± 0.005 (9.0)	0.840 ± 0.005 (2.5)	0.794 ± 0.000 (17.5)
25	0.771 ± 0.003 (12.5)	0.772 ± 0.004 (9.5)	0.771 ± 0.004 (12.5)	0.765 ± 0.004 (15.0)	0.773 ± 0.003 (7.5)	0.781 ± 0.003 (4.0)
26	0.868 ± 0.005 (1.5)	0.867 ± 0.004 (3.5)	0.864 ± 0.004 (9.0)	0.863 ± 0.004 (10.0)	0.865 ± 0.005 (7.5)	0.858 ± 0.005 (12.5)
27	1.000 ± 0.000 (5.0)	1.000 ± 0.000 (5.0)	1.000 ± 0.000 (5.0)	1.000 ± 0.000 (5.0)	1.000 ± 0.000 (5.0)	0.987 ± 0.002 (16.0)
28	0.628 ± 0.002 (9.0)	0.628 ± 0.002 (9.0)	0.627 ± 0.002 (14.5)	0.629 ± 0.002 (2.5)	0.627 ± 0.002 (14.5)	0.629 ± 0.002 (2.5)
29	0.638 ± 0.005 (11.5)	0.637 ± 0.004 (15.0)	0.637 ± 0.003 (15.0)	0.638 ± 0.005 (11.5)	0.635 ± 0.004 (20.0)	0.641 ± 0.004 (6.0)
30	0.625 ± 0.002	0.625 ± 0.002	0.624 ± 0.001	0.625 ± 0.001	0.624 ± 0.002	0.622 ± 0.001

	(4.5)	(4.5)	(9.5)	(4.5)	(9.5)	(15.0)
Rank	8.483	8.633	9.800	10.383	9.317	13.050

Table 4. Testing Accuracies of ELMs Corresponding to 24 Probability Distributions on 30 UCI Data Sets

	D ₁ :Beta	D ₂ :Cauchy	D ₃ :Central Chi	D ₄ :Chi	D ₅ :Erlang	D ₆ :Exponential
1	0.806 ± 0.009 (2.5)	0.790 ± 0.012 (23.0)	0.803 ± 0.010 (7.5)	0.804 ± 0.011 (5.5)	0.804 ± 0.011 (5.5)	0.802 ± 0.008 (9.0)
2	0.794 ± 0.004 (10.5)	0.790 ± 0.006 (23.5)	0.795 ± 0.005 (6.5)	0.795 ± 0.004 (6.5)	0.796 ± 0.005 (3.5)	0.792 ± 0.005 (17.5)
3	0.965 ± 0.003 (11.0)	0.964 ± 0.004 (16.5)	0.968 ± 0.004 (1.5)	0.965 ± 0.003 (11.0)	0.967 ± 0.003 (5.0)	0.966 ± 0.004 (8.0)
4	0.962 ± 0.006 (4.5)	0.934 ± 0.009 (12.0)	0.823 ± 0.009 (22.0)	0.879 ± 0.008 (17.0)	0.818 ± 0.010 (23.0)	0.913 ± 0.006 (16.0)
5	0.714 ± 0.022 (19.5)	0.726 ± 0.025 (18.0)	0.743 ± 0.009 (14.0)	0.729 ± 0.014 (17.0)	0.745 ± 0.007 (12.0)	0.704 ± 0.019 (22.5)
6	0.540 ± 0.012 (16.0)	0.535 ± 0.013 (18.0)	0.533 ± 0.017 (20.5)	0.545 ± 0.011 (8.5)	0.533 ± 0.014 (20.5)	0.540 ± 0.014 (16.0)
7	0.767 ± 0.005 (18.0)	0.772 ± 0.009 (8.5)	0.774 ± 0.007 (5.5)	0.771 ± 0.007 (10.0)	0.776 ± 0.008 (1.5)	0.772 ± 0.007 (8.5)
8	0.630 ± 0.014 (11.0)	0.636 ± 0.018 (9.0)	0.588 ± 0.007 (20.0)	0.607 ± 0.013 (17.0)	0.587 ± 0.006 (21.0)	0.612 ± 0.018 (15.5)
9	0.852 ± 0.009 (5.5)	0.844 ± 0.010 (19.0)	0.843 ± 0.010 (20.0)	0.850 ± 0.008 (10.5)	0.839 ± 0.008 (22.0)	0.853 ± 0.008 (3.5)
10	0.647 ± 0.018 (19.0)	0.644 ± 0.019 (22.0)	0.649 ± 0.023 (14.0)	0.657 ± 0.018 (1.0)	0.655 ± 0.021 (4.0)	0.649 ± 0.018 (14.0)
11	0.734 ± 0.008 (3.5)	0.709 ± 0.012 (22.0)	0.714 ± 0.012 (13.0)	0.723 ± 0.009 (10.0)	0.712 ± 0.012 (15.5)	0.715 ± 0.010 (11.5)
12	0.801 ± 0.018 (5.0)	0.798 ± 0.019 (10.5)	0.728 ± 0.021 (22.0)	0.748 ± 0.012 (18.0)	0.716 ± 0.015 (24.0)	0.779 ± 0.019 (13.0)
13	0.923 ± 0.003 (1.0)	0.903 ± 0.005 (16.0)	0.873 ± 0.006 (20.0)	0.902 ± 0.004 (17.0)	0.866 ± 0.006 (22.0)	0.912 ± 0.003 (15.0)
14	0.803 ± 0.012 (11.0)	0.840 ± 0.019 (10.0)	0.642 ± 0.001 (21.5)	0.649 ± 0.004 (18.0)	0.641 ± 0.000 (23.5)	0.713 ± 0.012 (13.0)
15	0.953 ± 0.010 (5.0)	0.941 ± 0.012 (11.0)	0.943 ± 0.012 (8.0)	0.945 ± 0.008 (7.0)	0.942 ± 0.009 (9.5)	0.942 ± 0.011 (9.5)
16	0.805 ± 0.012 (11.0)	0.837 ± 0.021 (10.0)	0.642 ± 0.002 (21.5)	0.649 ± 0.005 (19.0)	0.641 ± 0.000 (23.5)	0.710 ± 0.012 (13.0)
17	0.842 ± 0.003 (4.0)	0.835 ± 0.005 (19.0)	0.834 ± 0.004 (20.0)	0.838 ± 0.004 (10.0)	0.831 ± 0.004 (21.0)	0.842 ± 0.003 (4.0)
18	0.565 ± 0.000 (17.5)	0.695 ± 0.017 (10.0)	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (17.5)
19	0.939 ± 0.003 (22.5)	0.947 ± 0.004 (1.0)	0.941 ± 0.003 (10.0)	0.940 ± 0.003 (17.5)	0.942 ± 0.003 (4.5)	0.941 ± 0.003 (10.0)
20	0.932 ± 0.006 (24.0)	0.940 ± 0.006 (10.5)	0.944 ± 0.005 (2.0)	0.943 ± 0.005 (5.0)	0.944 ± 0.005 (2.0)	0.942 ± 0.004 (8.0)
21	0.903 ± 0.017 (1.0)	0.851 ± 0.019 (21.0)	0.855 ± 0.018 (19.0)	0.876 ± 0.018 (8.0)	0.853 ± 0.017 (20.0)	0.884 ± 0.020 (6.0)
22	0.767 ± 0.006 (1.0)	0.760 ± 0.007 (19.0)	0.756 ± 0.008 (21.0)	0.766 ± 0.006 (4.0)	0.757 ± 0.008 (20.0)	0.765 ± 0.007 (9.5)
23	0.698 ± 0.026 (10.0)	0.686 ± 0.031 (12.0)	0.534 ± 0.000 (22.5)	0.539 ± 0.005 (17.5)	0.534 ± 0.000 (22.5)	0.629 ± 0.017 (13.0)
24	0.790 ± 0.003 (14.0)	0.758 ± 0.010 (20.0)	0.794 ± 0.000 (7.0)	0.794 ± 0.000 (7.0)	0.794 ± 0.000 (7.0)	0.794 ± 0.000 (7.0)
25	0.734 ± 0.009 (1.0)	0.688 ± 0.012 (18.0)	0.662 ± 0.012 (20.5)	0.703 ± 0.009 (16.0)	0.662 ± 0.011 (20.5)	0.713 ± 0.010 (14.0)
26	0.783 ± 0.011 (1.0)	0.703 ± 0.021 (19.0)	0.679 ± 0.016 (20.0)	0.759 ± 0.012 (16.0)	0.673 ± 0.015 (21.0)	0.763 ± 0.013 (15.0)
27	0.971 ± 0.011 (9.0)	0.947 ± 0.017 (12.5)	0.847 ± 0.020 (20.0)	0.908 ± 0.018 (17.0)	0.835 ± 0.018 (21.0)	0.946 ± 0.014 (14.0)
28	0.598 ± 0.004 (7.5)	0.593 ± 0.008 (19.5)	0.592 ± 0.008 (21.0)	0.599 ± 0.007 (4.0)	0.593 ± 0.005 (19.5)	0.599 ± 0.006 (4.0)
29	0.527 ± 0.015 (22.0)	0.533 ± 0.015 (5.5)	0.532 ± 0.015 (10.0)	0.535 ± 0.013 (2.5)	0.532 ± 0.015 (10.0)	0.531 ± 0.013 (14.5)

	0.597 ± 0.004 (6.5)	0.595 ± 0.005 (14.0)	0.590 ± 0.004 (20.0)	0.594 ± 0.005 (17.0)	0.587 ± 0.004 (21.5)	0.596 ± 0.005 (11.5)
Rank	9.833	15.000	15.600	11.733	15.467	11.783
	D_7 :F	D_8 :Frechet	D_9 :Gamma	D_{10} :Gumbel	D_{11} :Hyperbolic	D_{12} :Laplace
1	0.789 ± 0.013 (24.0)	0.806 ± 0.011 (2.5)	0.803 ± 0.009 (7.5)	0.800 ± 0.010 (12.0)	0.796 ± 0.007 (19.0)	0.796 ± 0.011 (19.0)
2	0.795 ± 0.005 (6.5)	0.796 ± 0.004 (3.5)	0.795 ± 0.004 (6.5)	0.793 ± 0.005 (13.5)	0.792 ± 0.005 (17.5)	0.791 ± 0.005 (21.5)
3	0.968 ± 0.003 (1.5)	0.964 ± 0.003 (16.5)	0.965 ± 0.003 (11.0)	0.967 ± 0.004 (5.0)	0.964 ± 0.003 (16.5)	0.964 ± 0.003 (16.5)
4	0.709 ± 0.039 (24.0)	0.916 ± 0.007 (14.0)	0.917 ± 0.007 (13.0)	0.825 ± 0.011 (20.0)	0.962 ± 0.006 (4.5)	0.961 ± 0.005 (7.5)
5	0.755 ± 0.006 (9.0)	0.704 ± 0.018 (22.5)	0.704 ± 0.020 (22.5)	0.743 ± 0.012 (14.0)	0.760 ± 0.018 (5.0)	0.752 ± 0.025 (10.0)
6	0.528 ± 0.012 (24.0)	0.543 ± 0.012 (12.0)	0.545 ± 0.013 (8.5)	0.531 ± 0.013 (22.5)	0.545 ± 0.014 (8.5)	0.546 ± 0.013 (5.5)
7	0.775 ± 0.008 (3.0)	0.768 ± 0.006 (15.0)	0.769 ± 0.008 (12.0)	0.774 ± 0.008 (5.5)	0.765 ± 0.007 (21.0)	0.770 ± 0.008 (11.0)
8	0.585 ± 0.005 (23.0)	0.614 ± 0.013 (14.0)	0.612 ± 0.014 (15.5)	0.586 ± 0.008 (22.0)	0.647 ± 0.014 (3.5)	0.645 ± 0.019 (7.0)
9	0.836 ± 0.010 (24.0)	0.851 ± 0.007 (7.5)	0.853 ± 0.007 (3.5)	0.838 ± 0.009 (23.0)	0.849 ± 0.008 (13.5)	0.849 ± 0.010 (13.5)
10	0.643 ± 0.019 (23.0)	0.656 ± 0.017 (2.0)	0.650 ± 0.019 (9.5)	0.641 ± 0.023 (24.0)	0.649 ± 0.017 (14.0)	0.655 ± 0.017 (4.0)
11	0.710 ± 0.010 (19.5)	0.734 ± 0.007 (3.5)	0.735 ± 0.008 (2.0)	0.707 ± 0.011 (24.0)	0.712 ± 0.011 (15.5)	0.710 ± 0.014 (19.5)
12	0.729 ± 0.020 (21.0)	0.775 ± 0.016 (15.0)	0.778 ± 0.014 (14.0)	0.734 ± 0.016 (20.0)	0.803 ± 0.014 (2.5)	0.800 ± 0.017 (8.0)
13	0.797 ± 0.011 (24.0)	0.916 ± 0.003 (4.5)	0.914 ± 0.002 (12.5)	0.858 ± 0.006 (23.0)	0.915 ± 0.003 (8.5)	0.915 ± 0.004 (8.5)
14	0.642 ± 0.002 (21.5)	0.691 ± 0.010 (14.0)	0.687 ± 0.008 (15.0)	0.648 ± 0.005 (19.0)	0.884 ± 0.014 (2.0)	0.874 ± 0.015 (9.0)
15	0.937 ± 0.010 (14.0)	0.955 ± 0.008 (3.5)	0.955 ± 0.008 (3.5)	0.940 ± 0.010 (12.5)	0.935 ± 0.013 (21.0)	0.935 ± 0.011 (21.0)
16	0.642 ± 0.002 (21.5)	0.693 ± 0.009 (14.0)	0.686 ± 0.008 (15.0)	0.650 ± 0.005 (18.0)	0.881 ± 0.012 (5.0)	0.878 ± 0.013 (8.0)
17	0.824 ± 0.006 (24.0)	0.842 ± 0.004 (4.0)	0.842 ± 0.004 (4.0)	0.830 ± 0.005 (22.0)	0.838 ± 0.003 (10.0)	0.837 ± 0.004 (15.5)
18	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (17.5)	0.749 ± 0.018 (4.5)	0.735 ± 0.021 (9.0)
19	0.946 ± 0.004 (2.0)	0.939 ± 0.003 (22.5)	0.940 ± 0.003 (17.5)	0.942 ± 0.003 (4.5)	0.941 ± 0.003 (10.0)	0.940 ± 0.003 (17.5)
20	0.942 ± 0.005 (8.0)	0.940 ± 0.005 (10.5)	0.939 ± 0.004 (12.0)	0.943 ± 0.005 (5.0)	0.935 ± 0.005 (20.5)	0.935 ± 0.006 (20.5)
21	0.816 ± 0.019 (24.0)	0.894 ± 0.016 (5.0)	0.895 ± 0.016 (3.5)	0.846 ± 0.020 (23.0)	0.869 ± 0.015 (15.0)	0.872 ± 0.018 (10.5)
22	0.752 ± 0.009 (24.0)	0.765 ± 0.006 (9.5)	0.765 ± 0.006 (9.5)	0.755 ± 0.010 (22.5)	0.765 ± 0.008 (9.5)	0.763 ± 0.007 (15.5)
23	0.534 ± 0.000 (22.5)	0.592 ± 0.013 (14.0)	0.572 ± 0.009 (15.0)	0.537 ± 0.005 (19.0)	0.757 ± 0.025 (4.0)	0.748 ± 0.028 (8.0)
24	0.794 ± 0.000 (7.0)	0.794 ± 0.000 (7.0)	0.794 ± 0.000 (7.0)	0.794 ± 0.000 (7.0)	0.757 ± 0.016 (21.0)	0.749 ± 0.017 (24.0)
25	0.613 ± 0.016 (24.0)	0.721 ± 0.009 (3.5)	0.719 ± 0.011 (6.0)	0.648 ± 0.015 (23.0)	0.720 ± 0.010 (5.0)	0.716 ± 0.011 (8.5)
26	0.528 ± 0.020 (24.0)	0.776 ± 0.014 (12.5)	0.780 ± 0.013 (3.0)	0.631 ± 0.018 (23.0)	0.779 ± 0.013 (5.5)	0.777 ± 0.012 (10.5)
27	0.808 ± 0.027 (24.0)	0.944 ± 0.011 (15.0)	0.947 ± 0.015 (12.5)	0.832 ± 0.021 (22.5)	0.975 ± 0.012 (3.0)	0.975 ± 0.012 (3.0)
28	0.579 ± 0.008 (24.0)	0.599 ± 0.008 (4.0)	0.601 ± 0.006 (1.0)	0.588 ± 0.008 (23.0)	0.597 ± 0.007 (10.5)	0.597 ± 0.006 (10.5)
29	0.519 ± 0.015 (24.0)	0.533 ± 0.011 (5.5)	0.534 ± 0.013 (4.0)	0.527 ± 0.017 (22.0)	0.532 ± 0.014 (10.0)	0.530 ± 0.016 (16.5)
30	0.579 ± 0.005 (24.0)	0.595 ± 0.004 (14.0)	0.594 ± 0.004 (17.0)	0.587 ± 0.005 (21.5)	0.597 ± 0.004 (6.5)	0.598 ± 0.005 (1.5)

Rank	18.550	10.283	10.033	17.783	10.417	12.017
	D ₁₃ :Logistic	D ₁₄ :Lognormal	D ₁₅ :Normal	D ₁₆ :Pareto	D ₁₇ :Rayleigh	D ₁₈ :Semicircle
1	0.801 ± 0.009 (10.5)	0.799 ± 0.009 (13.0)	0.796 ± 0.010 (19.0)	0.801 ± 0.007 (10.5)	0.805 ± 0.010 (4.0)	0.798 ± 0.008 (14.0)
2	0.794 ± 0.005 (10.5)	0.793 ± 0.005 (13.5)	0.792 ± 0.004 (17.5)	0.797 ± 0.005 (2.0)	0.794 ± 0.005 (10.5)	0.792 ± 0.004 (17.5)
3	0.967 ± 0.004 (5.0)	0.967 ± 0.003 (5.0)	0.963 ± 0.004 (22.0)	0.967 ± 0.003 (5.0)	0.965 ± 0.003 (11.0)	0.963 ± 0.003 (22.0)
4	0.824 ± 0.009 (21.0)	0.857 ± 0.010 (18.0)	0.963 ± 0.004 (1.5)	0.854 ± 0.008 (19.0)	0.949 ± 0.008 (11.0)	0.960 ± 0.005 (9.5)
5	0.743 ± 0.011 (14.0)	0.733 ± 0.013 (16.0)	0.759 ± 0.019 (6.5)	0.748 ± 0.009 (11.0)	0.704 ± 0.018 (22.5)	0.762 ± 0.021 (3.5)
6	0.531 ± 0.012 (22.5)	0.534 ± 0.017 (19.0)	0.541 ± 0.013 (14.0)	0.542 ± 0.013 (13.0)	0.545 ± 0.015 (8.5)	0.550 ± 0.011 (1.5)
7	0.774 ± 0.008 (5.5)	0.776 ± 0.008 (1.5)	0.766 ± 0.008 (19.0)	0.774 ± 0.008 (5.5)	0.768 ± 0.007 (15.0)	0.768 ± 0.006 (15.0)
8	0.584 ± 0.008 (24.0)	0.596 ± 0.013 (18.0)	0.646 ± 0.013 (5.5)	0.589 ± 0.007 (19.0)	0.629 ± 0.014 (12.0)	0.643 ± 0.016 (8.0)
9	0.842 ± 0.010 (21.0)	0.847 ± 0.010 (16.0)	0.850 ± 0.011 (10.5)	0.854 ± 0.009 (2.0)	0.850 ± 0.008 (10.5)	0.847 ± 0.010 (16.0)
10	0.649 ± 0.023 (14.0)	0.649 ± 0.019 (14.0)	0.646 ± 0.016 (21.0)	0.649 ± 0.020 (14.0)	0.650 ± 0.023 (9.5)	0.649 ± 0.021 (14.0)
11	0.713 ± 0.010 (14.0)	0.715 ± 0.010 (11.5)	0.710 ± 0.009 (19.5)	0.729 ± 0.009 (6.0)	0.732 ± 0.011 (5.0)	0.725 ± 0.010 (7.0)
12	0.722 ± 0.017 (23.0)	0.753 ± 0.017 (17.0)	0.803 ± 0.017 (2.5)	0.740 ± 0.017 (19.0)	0.797 ± 0.016 (12.0)	0.801 ± 0.014 (5.0)
13	0.869 ± 0.006 (21.0)	0.894 ± 0.004 (18.0)	0.915 ± 0.003 (8.5)	0.892 ± 0.003 (19.0)	0.921 ± 0.003 (2.0)	0.914 ± 0.003 (12.5)
14	0.644 ± 0.004 (20.0)	0.655 ± 0.006 (17.0)	0.885 ± 0.014 (1.0)	0.641 ± 0.000 (23.5)	0.797 ± 0.010 (12.0)	0.879 ± 0.013 (5.5)
15	0.936 ± 0.011 (17.0)	0.940 ± 0.010 (12.5)	0.935 ± 0.011 (21.0)	0.958 ± 0.010 (2.0)	0.951 ± 0.009 (6.0)	0.936 ± 0.013 (17.0)
16	0.645 ± 0.004 (20.0)	0.655 ± 0.006 (17.0)	0.882 ± 0.012 (2.5)	0.641 ± 0.000 (23.5)	0.796 ± 0.012 (12.0)	0.881 ± 0.014 (5.0)
17	0.828 ± 0.005 (23.0)	0.839 ± 0.005 (7.0)	0.837 ± 0.004 (15.5)	0.837 ± 0.004 (15.5)	0.843 ± 0.003 (1.0)	0.838 ± 0.003 (10.0)
18	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (17.5)	0.746 ± 0.016 (6.0)	0.565 ± 0.000 (17.5)	0.565 ± 0.000 (17.5)	0.763 ± 0.018 (1.0)
19	0.941 ± 0.004 (10.0)	0.943 ± 0.003 (3.0)	0.941 ± 0.003 (10.0)	0.941 ± 0.004 (10.0)	0.939 ± 0.003 (22.5)	0.941 ± 0.004 (10.0)
20	0.943 ± 0.005 (5.0)	0.944 ± 0.005 (2.0)	0.936 ± 0.005 (16.0)	0.942 ± 0.005 (8.0)	0.936 ± 0.005 (16.0)	0.935 ± 0.007 (20.5)
21	0.850 ± 0.020 (22.0)	0.864 ± 0.017 (18.0)	0.868 ± 0.016 (16.5)	0.880 ± 0.016 (7.0)	0.896 ± 0.018 (2.0)	0.870 ± 0.016 (14.0)
22	0.755 ± 0.009 (22.5)	0.761 ± 0.007 (18.0)	0.766 ± 0.007 (4.0)	0.762 ± 0.005 (17.0)	0.766 ± 0.007 (4.0)	0.766 ± 0.007 (4.0)
23	0.536 ± 0.003 (20.0)	0.539 ± 0.006 (17.5)	0.754 ± 0.023 (7.0)	0.534 ± 0.000 (22.5)	0.689 ± 0.021 (11.0)	0.759 ± 0.025 (2.0)
24	0.794 ± 0.000 (7.0)	0.794 ± 0.000 (7.0)	0.759 ± 0.016 (19.0)	0.794 ± 0.000 (7.0)	0.794 ± 0.001 (7.0)	0.773 ± 0.015 (15.5)
25	0.650 ± 0.012 (22.0)	0.687 ± 0.013 (19.0)	0.716 ± 0.010 (8.5)	0.696 ± 0.009 (17.0)	0.727 ± 0.008 (2.0)	0.714 ± 0.010 (13.0)
26	0.658 ± 0.018 (22.0)	0.724 ± 0.015 (18.0)	0.778 ± 0.013 (8.0)	0.755 ± 0.014 (17.0)	0.778 ± 0.011 (8.0)	0.777 ± 0.011 (10.5)
27	0.832 ± 0.022 (22.5)	0.889 ± 0.016 (18.0)	0.975 ± 0.011 (3.0)	0.887 ± 0.018 (19.0)	0.969 ± 0.011 (11.0)	0.972 ± 0.009 (7.5)
28	0.591 ± 0.007 (22.0)	0.594 ± 0.007 (17.5)	0.597 ± 0.007 (10.5)	0.596 ± 0.006 (14.5)	0.599 ± 0.005 (4.0)	0.594 ± 0.006 (17.5)
29	0.528 ± 0.014 (20.0)	0.535 ± 0.012 (2.5)	0.529 ± 0.009 (18.5)	0.541 ± 0.015 (1.0)	0.527 ± 0.014 (22.0)	0.530 ± 0.011 (16.5)
30	0.586 ± 0.005 (23.0)	0.595 ± 0.005 (14.0)	0.597 ± 0.005 (6.5)	0.592 ± 0.005 (19.0)	0.596 ± 0.005 (11.5)	0.597 ± 0.004 (6.5)
Rank	17.383	13.533	11.350	12.867	10.100	10.717
	D ₁₉ :Student's-t	D ₂₀ :Triangular	D ₂₁ :Truncated	D ₂₂ :Uniform	D ₂₃ :Von Mises	D ₂₄ :Weibull

1	0.795 ± 0.011 (22.0)	0.796 ± 0.011 (19.0)	0.797 ± 0.010 (15.5)	0.796 ± 0.010 (19.0)	0.797 ± 0.012 (15.5)	0.808 ± 0.008 (1.0)
2	0.791 ± 0.004 (21.5)	0.792 ± 0.004 (17.5)	0.794 ± 0.005 (10.5)	0.790 ± 0.005 (23.5)	0.792 ± 0.004 (17.5)	0.798 ± 0.003 (1.0)
3	0.964 ± 0.003 (16.5)	0.965 ± 0.003 (11.0)	0.963 ± 0.003 (22.0)	0.964 ± 0.003 (16.5)	0.963 ± 0.003 (22.0)	0.963 ± 0.002 (22.0)
4	0.962 ± 0.005 (4.5)	0.963 ± 0.003 (1.5)	0.960 ± 0.005 (9.5)	0.962 ± 0.006 (4.5)	0.961 ± 0.004 (7.5)	0.914 ± 0.007 (15.0)
5	0.762 ± 0.021 (3.5)	0.759 ± 0.020 (6.5)	0.764 ± 0.017 (2.0)	0.758 ± 0.019 (8.0)	0.772 ± 0.022 (1.0)	0.714 ± 0.015 (19.5)
6	0.544 ± 0.014 (11.0)	0.547 ± 0.013 (4.0)	0.549 ± 0.014 (3.0)	0.540 ± 0.014 (16.0)	0.550 ± 0.013 (1.5)	0.546 ± 0.014 (5.5)
7	0.765 ± 0.006 (21.0)	0.765 ± 0.007 (21.0)	0.764 ± 0.006 (23.0)	0.763 ± 0.009 (24.0)	0.768 ± 0.008 (15.0)	0.768 ± 0.006 (15.0)
8	0.647 ± 0.018 (3.5)	0.646 ± 0.017 (5.5)	0.648 ± 0.015 (2.0)	0.633 ± 0.013 (10.0)	0.649 ± 0.017 (1.0)	0.616 ± 0.015 (13.0)
9	0.851 ± 0.009 (7.5)	0.850 ± 0.008 (10.5)	0.846 ± 0.008 (18.0)	0.852 ± 0.007 (5.5)	0.847 ± 0.006 (16.0)	0.857 ± 0.005 (1.0)
10	0.655 ± 0.019 (4.0)	0.653 ± 0.018 (6.5)	0.647 ± 0.018 (19.0)	0.652 ± 0.019 (8.0)	0.647 ± 0.016 (19.0)	0.653 ± 0.020 (6.5)
11	0.710 ± 0.011 (19.5)	0.711 ± 0.011 (17.0)	0.724 ± 0.010 (8.5)	0.708 ± 0.011 (23.0)	0.724 ± 0.013 (8.5)	0.741 ± 0.008 (1.0)
12	0.800 ± 0.011 (8.0)	0.801 ± 0.013 (5.0)	0.798 ± 0.016 (10.5)	0.805 ± 0.017 (1.0)	0.800 ± 0.014 (8.0)	0.772 ± 0.017 (16.0)
13	0.916 ± 0.003 (4.5)	0.916 ± 0.003 (4.5)	0.914 ± 0.003 (12.5)	0.916 ± 0.003 (4.5)	0.915 ± 0.003 (8.5)	0.914 ± 0.003 (12.5)
14	0.876 ± 0.015 (7.0)	0.881 ± 0.012 (3.0)	0.879 ± 0.012 (5.5)	0.875 ± 0.012 (8.0)	0.880 ± 0.013 (4.0)	0.676 ± 0.007 (16.0)
15	0.936 ± 0.011 (17.0)	0.936 ± 0.013 (17.0)	0.934 ± 0.013 (23.5)	0.936 ± 0.011 (17.0)	0.934 ± 0.014 (23.5)	0.965 ± 0.008 (1.0)
16	0.880 ± 0.016 (7.0)	0.883 ± 0.011 (1.0)	0.881 ± 0.011 (5.0)	0.877 ± 0.013 (9.0)	0.882 ± 0.011 (2.5)	0.678 ± 0.006 (16.0)
17	0.838 ± 0.004 (10.0)	0.837 ± 0.004 (15.5)	0.838 ± 0.004 (10.0)	0.837 ± 0.003 (15.5)	0.837 ± 0.003 (15.5)	0.842 ± 0.003 (4.0)
18	0.738 ± 0.022 (7.5)	0.749 ± 0.018 (4.5)	0.759 ± 0.015 (2.0)	0.738 ± 0.020 (7.5)	0.757 ± 0.016 (3.0)	0.565 ± 0.000 (17.5)
19	0.940 ± 0.003 (17.5)	0.940 ± 0.003 (17.5)	0.941 ± 0.003 (10.0)	0.940 ± 0.003 (17.5)	0.941 ± 0.003 (10.0)	0.939 ± 0.003 (22.5)
20	0.937 ± 0.006 (13.5)	0.936 ± 0.005 (16.0)	0.935 ± 0.006 (20.5)	0.935 ± 0.006 (20.5)	0.935 ± 0.006 (20.5)	0.937 ± 0.005 (13.5)
21	0.871 ± 0.017 (12.5)	0.871 ± 0.016 (12.5)	0.872 ± 0.017 (10.5)	0.868 ± 0.017 (16.5)	0.875 ± 0.016 (9.0)	0.895 ± 0.019 (3.5)
22	0.763 ± 0.007 (15.5)	0.765 ± 0.008 (9.5)	0.764 ± 0.008 (13.5)	0.764 ± 0.007 (13.5)	0.765 ± 0.008 (9.5)	0.766 ± 0.006 (4.0)
23	0.756 ± 0.023 (5.0)	0.755 ± 0.030 (6.0)	0.766 ± 0.024 (1.0)	0.746 ± 0.026 (9.0)	0.758 ± 0.027 (3.0)	0.563 ± 0.002 (16.0)
24	0.754 ± 0.017 (22.0)	0.761 ± 0.020 (18.0)	0.773 ± 0.017 (15.5)	0.753 ± 0.015 (23.0)	0.768 ± 0.016 (17.0)	0.794 ± 0.000 (7.0)
25	0.715 ± 0.011 (11.0)	0.715 ± 0.010 (11.0)	0.715 ± 0.009 (11.0)	0.706 ± 0.008 (15.0)	0.718 ± 0.011 (7.0)	0.721 ± 0.007 (3.5)
26	0.780 ± 0.014 (3.0)	0.780 ± 0.012 (3.0)	0.778 ± 0.013 (8.0)	0.774 ± 0.013 (14.0)	0.779 ± 0.013 (5.5)	0.776 ± 0.011 (12.5)
27	0.974 ± 0.009 (5.5)	0.974 ± 0.011 (5.5)	0.977 ± 0.009 (1.0)	0.972 ± 0.010 (7.5)	0.970 ± 0.011 (10.0)	0.940 ± 0.012 (16.0)
28	0.597 ± 0.007 (10.5)	0.596 ± 0.006 (14.5)	0.596 ± 0.005 (14.5)	0.598 ± 0.006 (7.5)	0.596 ± 0.007 (14.5)	0.599 ± 0.006 (4.0)
29	0.529 ± 0.015 (18.5)	0.532 ± 0.016 (10.0)	0.532 ± 0.015 (10.0)	0.532 ± 0.012 (10.0)	0.532 ± 0.012 (10.0)	0.531 ± 0.013 (14.5)
30	0.597 ± 0.005 (6.5)	0.597 ± 0.005 (6.5)	0.597 ± 0.005 (6.5)	0.597 ± 0.005 (6.5)	0.598 ± 0.005 (1.5)	0.594 ± 0.004 (17.0)
Rank	11.217	10.017	10.800	12.700	10.233	10.583

is the hidden-layer output matrix for testing instances, $g(z) = \frac{1}{1+e^{-z}}$ is sigmoid activation function,

$$W_{D \times L} = [w_1 \quad w_2 \quad \cdots \quad w_L] = \begin{bmatrix} w_{11} & w_{21} & \cdots & w_{L1} \\ w_{12} & w_{22} & \cdots & w_{L2} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1D} & w_{2D} & \cdots & w_{LD} \end{bmatrix} \text{ and } b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_L \end{bmatrix} \quad (4)$$

are input-layer weight and hidden-layer biases which are usually determined by assigning the random numbers obeying uniform distribution. Our main work in this paper is to use the probability distributions in Table 1 to initialize $W_{D \times L}$ and b and compare the classification performances of corresponding ELMs.

3. Impact of Random Initialization on ELM's Generalization Performance

3.1. Data Preparation Guideline

In this experimental study, 30 UCI data sets are used to test the classification performances of ELMs corresponding to different random initializations. The detailed description of 30 data sets is summarized in Table 2. In order to use these data sets more efficiently and specifically, we preprocess them according to the following guidelines: (1) Delete nominal-value attributes. ELMs are mainly used to handle the classification problem with continuous-value attributes. Up to now, we don't find the study regarding how to use ELM to deal with nominal-value attributes. (2) Replace missing attribute-values. The unsupervised filter named *ReplaceMissingValues* in Weka [18] is used to replace all missing attribute-values in each data set. (3) Reduce large data sets. The unsupervised filter named *Resample* with the *sampleSizePercent* 10 in Weka is used to randomly reduce the sizes of three large data sets: Magic Telescope, Page Blocks and Wine Quality-White.

3.2. Experimental Setup

We use 10-fold cross-validation to get the training and testing accuracies corresponding to 24 ELMs (i.e., initializing the input-layer weights and hidden-layer biases with random numbers obeying 24 different probability distributions in Table 1 respectively) on every data set. In order to make the experimental results more reliable, 10-fold cross-validation on each data set is repeated 100 times for every ELM and the average values of 100 training and testing accuracies are used as the finally experimental results. In every time of 10-fold cross-validation, the different ELMs are trained on the same training sets and evaluated on the same testing set and the random initializations to input-layer weights and hidden-layer biases are independent.

The number of hidden-layer nodes in all ELMs used in our experimental study is 50. The necessary parameters in 24 probability distributions are initialized as follows: (1) $\alpha=2$ and $\beta=2$ for Beta; (2) $\lambda=1$ for Cauchy; (3) $k=2$ for Central Chi-Squared; (4) $k=2$ for Chi; (5) $\lambda=1$ and $k=2$ for Erlang; (6) $\lambda=1$ for Exponential; (7) $n_1=5$ and $n_2=2$ for F; (8) $s=1$ and $\alpha=3$ Frechet; (9) $k=9$ and $\theta=0.1$ Gamma; (10) $\mu=1$ and $\beta=2$ Gumbel; (11) no parameter for Hyperbolic Secant; (12) $\mu=0$ and $b=1$ for Laplace; (13) $\mu=9$ and $s=3$ for Logistic; (14) $\mu=0$ and $\sigma=1$ Lognormal; (15) $\mu=0$ and $\sigma=1$ Normal; (16) $x_m=1$ and $\alpha=3$ Pareto; (17) $\sigma=0.5$ for Rayleigh; (18) $r=1$ for Semicircle; (19) $\nu=5$ for Student's-t; (20) $a=-2.5$, $c=0$ and $b=2.5$ for Triangular; (21) $a=-1$, $b=1$, $\mu=0$ and $\sigma=1$ for Truncated Normal; (22) $a=-2.5$ and $b=2.5$ for Uniform; (23) $\mu=0$ and $k=4$ for Von Mises; and (24) $\lambda=1$ and $k=5$ for Weibull.

3.3. Experimental Result and Analysis

The training and testing accuracies of 24 ELMs on 30 UCI data sets are summarized in Tables 3 and 4 respectively. Here, we use Wilcoxon signed-ranks test and Friedman test [3] to conduct the statistical analysis to above-mentioned experimental results. The former is to assess whether the classification

accuracies of two ELMs differ on a single data set, while the latter is to test whether or not the classification accuracies of these 24 ELMs are equal on all data sets. Here, we only give the detailed procedures of these two statistical tests based on training accuracies of different ELMs. Similarly, the testing accuracies can be dealt with.

The compared objects of Wilcoxon signed-ranks test are the respective 100 training accuracies of two ELMs (ELM₁ and ELM₂) on a single data set.

1. Let $a_i^{(1)}$ and $a_i^{(2)}$ denote the training accuracies of ELM₁ and ELM₂ in the i -th ($i=1,2,\dots,100$) run of 10-fold cross-validation respectively.
2. $d_i = a_i^{(1)} - a_i^{(2)}, i=1,2,\dots,100$ denotes the difference between the training accuracies of ELM₁ and ELM₂ in the i -th run of 10-fold cross-validation.
3. Increasingly order $|d_i|, i=1,2,\dots,100$ starting with the smallest as 1. For d_i with same absolute values, the average of their ranks is assigned as the new rank. Let $rank(d_i)$ denote the rank of d_i .
4. Let R_+ and R_- denote the sums of ranks for the runs on which ELM₁ is better and worse than ELM₂ respectively. R_+ and R_- can be calculated as follows:

$$R_+ = \sum_{\substack{i=1 \\ d_i > 0}}^{100} rank(d_i) + \frac{1}{2} \sum_{\substack{i=1 \\ d_i = 0}}^{100} rank(d_i) \tag{5}$$

and

$$R_- = \sum_{\substack{i=1 \\ d_i < 0}}^{100} rank(d_i) + \frac{1}{2} \sum_{\substack{i=1 \\ d_i = 0}}^{100} rank(d_i). \tag{6}$$

5. Construct the statistic $z = \frac{\min(R_+, R_-) - \frac{1}{4}m(m+1)}{\sqrt{\frac{1}{24}m(m+1)(2m+1)}}$ based on the null-hypothesis, which indicates that ELM1

and ELM2 have the equal training accuracy on a given data set, m is the times of 10-fold cross-validations. In our study, $m=100$. When $m \geq 30$, the distribution of z is approximately normal.

6. Set the significance level $\alpha = 0.1$. If $z > z_{\frac{\alpha}{2}}$ or $z < -z_{\frac{\alpha}{2}}$, the null-hypothesis can be rejected, i.e., there are some differences between ELM₁ and ELM₂; else, we accept the null-hypothesis. The value $z_{\frac{0.1}{2}} = z_{0.05} = 1.645$ can be found from any statistical textbook.

After we compare the training accuracies of two ELMs on a single data set, Friedman test is used to analyze the difference of training accuracies corresponding to 24 ELMs on 30 data sets. The analyzed object is the experimental results in Table 3.

1. Increasingly order the average training accuracies of 24 ELMs on every data set. Then, the average value of ranks is assigned to the same training accuracy (This depends on the results obtained by previous Wilcoxon signed-ranks test: if the null-hypothesis is accepted, we say the training accuracies of two ELMs are same and thus have the same rank). In Table 3, the number in parentheses is the rank of training accuracy.
2. For 24 ELMs, the average value of ranks $R_i, i=1,2,\dots,24$ on 30 data sets is calculated respectively (i.e., the last line of Table 3).
3. Construct the statistic $F_F = \frac{(m-1)\chi_F^2}{m(k-1)-\chi_F^2}$ based on the null-hypothesis which indicates that all models have

the equal training accuracy on all data sets, where $\chi_F^2 = \frac{12m}{k(k+1)} \left[\sum_{j=1}^k R_j^2 - \frac{k(k+1)^2}{4} \right]$ is the Friedman statistic, k is

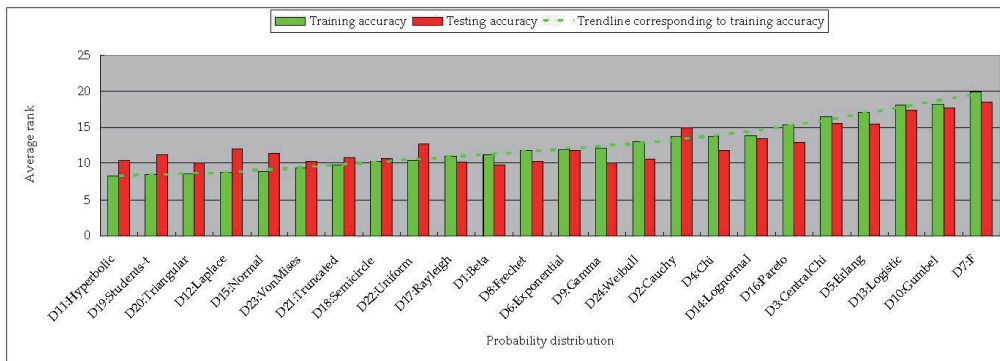
the number of learning models, m is the number of data sets. In this study, $k=24$ and $m=30$.

4. The statistic F_F obeys the F-distribution with $(k-1)$ and $(k-1) \times (m-1)$ degrees of freedom. Under the significance level $\alpha = 0.1$, we can obtain

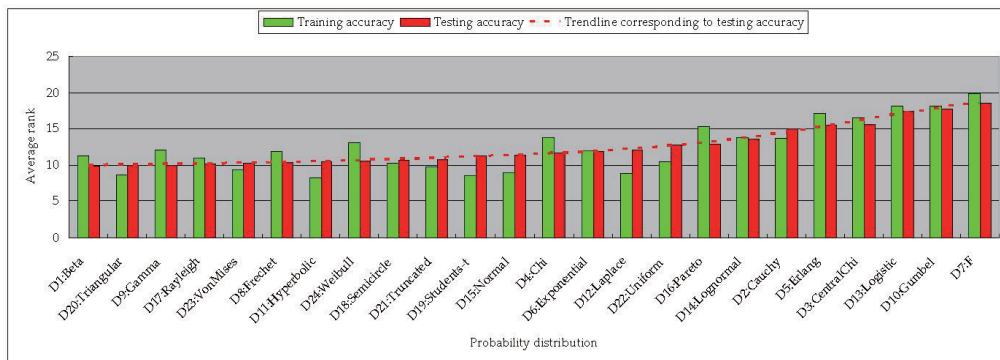
$$F_{\alpha}((k-1), (k-1) \times (m-1)) = F_{0.1}(23, 667) = 1.403. \tag{7}$$

5. If $F_F > F_{0.1}(23, 667)$, the null-hypothesis can be rejected. That is to say the training accuracies of these 24 ELMs are different on the 30 data sets.

Based on the analyzed results with Wilcoxon signed-ranks test and Friedman test, we increasingly order these 24 probability distributions in Fig. 1 according to the average ranks of training and testing accuracies respectively. From Fig. 1, we give the following interestingly experimental observations and usefully empirical conclusions.

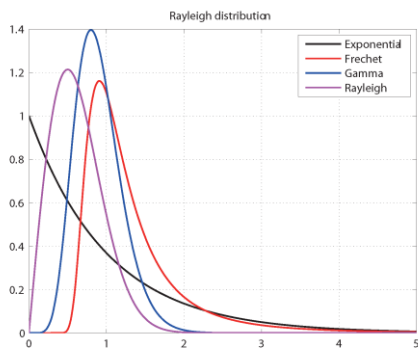


(a). Training Accuracy

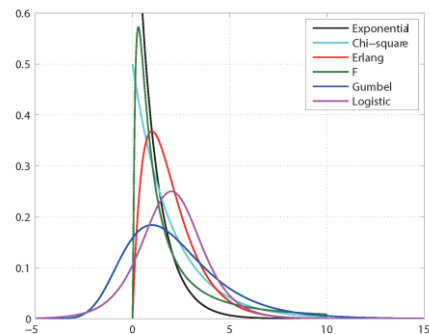


(b). Testing Accuracy

Fig. 1. Increasingly order these 24 probability distributions according to the average ranks of training and testing.



(a) Heavy-tailed distributions



(b) Light-tailed distributions

Fig. 2. Heavy-tailed and light-tailed distributions for ELM initialization.

1. ELMs initialized with *Hyperbolic Secant*, *Student's-t*, *Laplace* and *Normal* distributions obtain the better training accuracies. Meanwhile, these four distributions are more easily to cause the over-fitting of ELM. We can see that the orders of training accuracies corresponding to these four distributions are 1, 2, 4, and 5 respectively in Fig. 1-(a), while their orders of testing accuracies in Fig. 1-(b) are 7, 11, 15 and 12. That is to say *Hyperbolic Secant*, *Student's-t*, *Laplace* and *Normal* distributions make ELM obtain higher training accuracy and lower testing accuracies. Some common characteristics of these probability distributions are (1) their probability density functions $p(x)$ are bell-shaped; (2) all $p(x)$ are symmetrical with respect to $x=0$; and (3) the support intervals of $p(x)$ are $x \in (-\infty, +\infty)$.
2. ELMs initialized with *Gamma*, *Rayleigh* and *Frechet* distributions obtain the better testing accuracies. We can see that the orders of testing accuracies corresponding to these three distributions are 3, 4, and 6 respectively in Fig. 1-(b), while their orders of training accuracies in Fig. 1-(a) are 14, 10 and 12. This indicates that *Gamma*, *Rayleigh* and *Frechet* distributions make ELM have the better predictive capability to unseen instances. The common characteristics of these probability distributions include (1) their probability density functions $p(x)$ have the heavy tails; (2) $p(x)$ are asymmetrical with respect to $x=0$; and (3) the support intervals of $p(x)$ are $x \in (0, +\infty)$.
3. The mostly-used *Uniform* distribution cannot bring about the better testing accuracy for ELM. The order of testing accuracy of *Uniform* distribution is 16 in Fig. 1-(b). The distributions with light tail, e.g., central *Chi-Squared*, *Erlang*, *F*, *Gumbel* and *Logistic* are unsuited to initialize the input-layer weights and hidden-layer biases for ELM, because they lead to the worse training and testing accuracies on selected 30 data sets.

In Fig. 2, we plot the probability density functions for heavy-tailed *Gamma*, *Rayleigh* and *Frechet* in Fig. 2-(a) and light-tailed (*Central Chi-Squared*, *Erlang*, *F*, *Gumbel* and *Logistic* in Fig. 2-(b)) distributions. An intuitive explanation to heavy-tailed distribution is its probability density function $p_1(x)$ has a heavier tail than *Exponential* distribution, i.e., the right tail of distribution $p_1(x)$ will be located below *Exponential* distribution with the increase of x . In contrast with heavy-tailed distribution, the light-tailed distribution $p_2(x)$ has a lighter tail than *Exponential* distribution, the right tail of distribution $p_2(x)$ will be located above *Exponential* distribution with the increase of x . From Fig. 1, we can see that the heavy-tailed distributions (e.g., (*Gamma*, *Rayleigh* and *Frechet*)) obtain the better generalization performances than *Uniform* distribution and the light-tailed distributions (e.g., *Central Chi-Squared*, *Erlang*, *F*, *Gumbel* and *Logistic*) get the worse training and testing accuracies than *Uniform* distribution.

4. Conclusion

In this paper, we experimentally validate the impact of different continuous probability distributions on the classification accuracies of ELMs and give some important and useful enlightenments regarding ELM initialization. By initializing the input-layer weights and hidden-layer biases with random numbers obeying 24 different continuous probability distributions, we compare the training and testing accuracies of 24 ELMs on 30 UCI data sets. The experimental results and statistical analysis reflect that (1) the symmetrical and bell-shaped probability distributions have the higher training accuracies and meanwhile easily cause the over-fitting of ELM; (2) the heavy-tailed distributions bring about the higher testing accuracies for ELM than *Uniform* distribution; and (3) the light-tailed distributions obtain the worse training and testing accuracies than *Uniform* distribution.

Acknowledgement

We thank Editor-in-Chief and anonymous reviewers for their valuable comments and suggestions which help us to improve this paper significantly. This paper was supported by National Natural Science

Foundation of China (61503252) and China Postdoctoral Science Foundation (2016T90799).

References

- [1] Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing*, 70(1), 489-501.
- [2] Serre, D. (2006). *Matrices: Theory and Applications*. New York: Springer.
- [3] Huang, G. B., Chen, L., & Siew C. K. (2006). Universal approximation using incremental constructive feedforward networks with random hidden nodes. *IEEE Transactions on Neural Networks*, 17(4), 879-892.
- [4] Huang, G. B., Wang, D. H., & Lan, Y. (2011). Extreme learning machines: A survey. *International Journal of Machine Learning and Cybernetics*, 2(2), 107-122.
- [5] Zhu, Q. Y., Qin, A., Suganthan, P., & Huang, G. B. (2005). Evolutionary extreme learning machine. *Pattern Recognition*, 38(10), 1759-1763.
- [6] Cao, J., Lin, Z., & Huang, G. B. (2012). Self-adaptive evolutionary extreme learning machine. *Neural Processing Letters*, 36(3), 285-305.
- [7] Matias, T., Souza, F., Araújo, R., & Antunes, C. H. (2014). Learning of a single-hidden layer feedforward neural network using an optimized extreme learning machine. *Neurocomputing*, 129, 428-436.
- [8] Heeswijk, M. V., & Miche, Y. (2015). Binary/ternary extreme learning machines. *Neurocomputing*, 149, 187-197.
- [9] Balasundaram, S., & Gupta, D. (2014). On optimization based extreme learning machine in primal for regression and classification by functional iterative method. *International Journal of Machine Learning and Cybernetics*.
- [10] Deng, W., Zheng, Q., & Chen, L. (2009). Regularized extreme learning machine. *Proceedings of 2009 IEEE Symposium on Computational Intelligence and Data Mining* (PP. 389-395).
- [11] Horata, P., Chiewchanwattana, S., & Sunat, K. (2013). Robust extreme learning machine. *Neurocomputing*, 102, 31-44.
- [12] Huang, G. B., Li, M. B., Chen, L., & Siew, C. K. (2008). Incremental extreme learning machine with fully complex hidden nodes. *Neurocomputing*, 71(4-6), 576-583.
- [13] Miche, Y., Sorjamaa, A., Bas, P., Simula, O., Jutten, C., & Lendasse, A. (2010). OP-ELM: Optimally pruned extreme learning machine. *IEEE Transactions on Neural Networks*, 21(1), 158-162.
- [14] Wang, X. Z., Shao, Q. Y., Miao, Q., & Zhai, J. H. (2013). Architecture selection for networks trained with extreme learning machine using localized generalization error model. *Neurocomputing*, 102, 3-9.
- [15] Zhang, K., & Luo, M. X. (2015). Outlier-robust extreme learning machine for regression problems. *Neurocomputing*, 151, 1519-1527.
- [16] He, Y. L., Liu, J. N. K., Wang, X. Z., & Hu, Y. X. (2012). Optimal bandwidth selection for resubstitution entropy estimation. *Applied Mathematics and Computation*, 219(8), 3425-3460.
- [17] Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, 7(1), 1-30.
- [18] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: An update. *SIGKDD Explorations*, 11(1), 10-18.



Xiaofang Zhang completed her bachelor and master degree in computer science from Wuhan University of Science and Technology, Wuhan China. From 2008 to 2011, she worked as an assistant in Fujian Normal University Minnan Science and Technology Institute. From 2011 to 2016, she worked as a lecturer in Beijing Jiaotong University Haibin College. Since

2017, she is working as a lecturer in City College, Wuhan University of Science and Technology. Her main research interests include machine learning and big data mining.



Xiaoli Lin received the B.S and M.S degree from Wuhan University of Science and Technology of China, Hubei, China, in 2004 and 2007. She is currently a doctoral student in the College of Computer Science and Technology, Wuhan University of Science and Technology, China. She is mainly interested in computational biology and bioinformatics.



Rana Aamir Raza Ashfaq completed his Ph.D from Shenzhen University, Shenzhen China. He received his master degree in computer science from Blekinge Tekniska Hgskola (BTH), Sweden. He also received his bachelor and master degrees in computer science from Bahauddin Zakariya University, Multan, Pakistan. Since 2010 he is working as assistant professor in Department of Computer Science, Bahauddin Zakariya University, Multan, Pakistan. His main research interests include machine learning and big data mining.