Learning Accurate Representation to Nonstandard Tensors via a Mode-Aware Tucker Network

Hao Wu, *Member, IEEE*, Qu Wang, Xin Luo, *Fellow, IEEE*, and Zidong Wang, *Fellow, IEEE*

I. Introduction

THIS is the supplementary file for paper entitled *Learning Accurate Representation to Nonstandard Tensors via a Mode-Aware Tucker Network*. Supplementary equations and experimental results are put into this file.

II. Supplementary Proof of Convergence

First, we review the formula for MSGD and rewrite it to separate the deterministic gradient part from the random noise part.

$$\begin{aligned} & v_{n+1} = \gamma v_n + \eta \nabla_{\omega_n} J\left(\omega_n\right) + \eta\left(\nabla_{\omega_n} J\left(\omega_n, \xi_n\right) - \nabla_{\omega_n} J\left(\omega_n\right)\right), \\ & \omega_{n+1} = \omega_n - v_{n+1}. \end{aligned} \tag{S1}$$

This transformation borrows the idea of random approximation. Then calculate the following Taylor expansion and do some processing to get:

$$J\left(\omega_{n+1}\right) - J\left(\omega_n\right) = -\nabla_\omega J(\omega_n)^T v_n + \tfrac{1}{2} v_n^T H_{\omega\omega}\left(\zeta_n\right) v_n. \tag{S2}$$

Using the above Taylor formula, we can make the following calculation:

$$\begin{aligned} & \nabla_{\omega_n} J(\omega_n)^T v_n \\ & = \left(\nabla_{\omega_{n-1}-v_{n-1}} J\left(\omega_{n-1} - v_{n-1}\right)\right)^T \left(\gamma v_{n-1} + \eta_n \nabla_{\omega_n} J\left(\omega_n, \xi_n\right)\right) \\ & = \gamma\left(\nabla_{\omega_{n-1}} J\left(\omega_{n-1}\right) - H_{\omega\omega}\left(\zeta_{n-1}\right) v_{n-1}\right)^T v_{n-1} + \eta_n \nabla_{\omega_n} J\left(\omega_n\right) \nabla_{\omega_n} J\left(\omega_n, \xi_n\right) \\ & = \gamma \nabla_{\omega_{n-1}} J(\omega_{n-1})^T v_{n-1} - \gamma v_{n-1}^T H_{\omega\omega}\left(\zeta_{n-1}\right) v_{n-1} + \eta_n \nabla_{\omega_n} J\left(\omega_n\right) \nabla_{\omega_n} J\left(\omega_n, \xi_n\right). \end{aligned} \tag{S3}$$

Next, processing the above recursive formula, we get:

$$\nabla_{\omega_n} J(\omega_n)^T v_n = \gamma \nabla_{\omega_1} J(\omega_1)^T v_1 - \sum_{i=1}^{n-1} \gamma v_i^T H_{\omega\omega}\left(\zeta_i\right) v_i + \sum_{i=2}^{t} \gamma \eta_i \nabla_{\omega_i} J\left(\omega_i\right) \nabla_{\omega_i} J\left(\omega_i, \zeta_i\right). \tag{S4}$$

Then, bring (S4) back to (S2),

$$J\left(\omega_{n+1}\right) - J\left(\omega_n\right) = -\gamma \nabla_{\omega_1} J(\omega_1)^T v_1 - \sum_{i=2}^{t} \gamma \eta_i \nabla_{\omega_i} J\left(\omega_i\right) \nabla_{\omega_i} J\left(\omega_i, \xi_i\right) + \sum_{i=1}^{n-1} \gamma v_i^T H_{\omega\omega}\left(\zeta_i\right) v_i + \tfrac{1}{2} v_n^T H_{\omega\omega}\left(\zeta_n\right) v_n. \tag{S5}$$

Next, we will show that $J(\omega_n + 1)$ is almost certainly convergent. We do this by recursing on (S5) and dividing it into three parts:

$$\begin{aligned}
J\left(\omega_{t+1}\right) = & \underbrace{J\left(\omega_1\right) - \frac{\gamma-\gamma^{t+1}}{1-\gamma} \nabla_{\omega_1} J(\omega_1)^T v_1 + \frac{1-\gamma^t}{1-\gamma} \eta_1 \nabla_{\omega_1} J\left(\omega_1\right) \nabla_{\omega_1} J\left(\omega_1, \xi_1\right)}_{A} \\
& \underbrace{- \sum_{n=1}^{t} \frac{1-\gamma^{t-n+1}}{1-\gamma} \eta_n \nabla_{\omega_n} J(\omega_n)^T \nabla_{\omega_n} J\left(\omega_n, \xi_n\right)}_{B} \\
& \underbrace{- \frac{1}{2} \sum_{n=1}^{t} v_n^T H_{\omega\omega}\left(\zeta_n\right) v_n + \sum_{n=1}^{t} \frac{1-\gamma^{t-n+1}}{1-\gamma} v_n^T H_{\omega\omega}\left(\zeta_n\right) v_n}_{C} .
\end{aligned} \tag{S6}$$

Because $a < 1$, an is convergent, which ensures that part A is converged. Then, Lemma 3 ensures that part C converges almost everywhere. For Part B, we have the following proof:

$$\begin{aligned}
B = & \sum_{n=1}^{t} \tfrac{1-\gamma^{t-n+1}}{1-\gamma} \eta_n \nabla_{\omega_n} J(\omega_n)^T \nabla_{\omega_n} J\left(\omega_n, \xi_n\right) \\
= & \sum_{n=1}^{t} \tfrac{1-\gamma^{t-n+1}}{1-\gamma} \eta_n \|\nabla_{\omega_n} J\left(\omega_n\right)\|^2 \\
+ & \sum_{n=1}^{t} \tfrac{1-\gamma^{t-n+1}}{1-\gamma} \eta_n \nabla_{\omega_n} J(\omega_n)^T \left(\nabla_{\omega_n} J\left(\omega_n, \xi_n\right) - \nabla_{\omega_n} J\left(\omega_n, \xi_n\right)\right).
\end{aligned} \tag{S7}$$

By Lemma 4, we have:

$$\sum_{n=1}^{t} \tfrac{1-\gamma^{t-n+1}}{1-\gamma} \eta_n \|\nabla_{\omega_n} J\left(\omega_n\right)\|^2 < +\infty. \tag{S8}$$

According to Lemma 1 and Lemma 4, we can conclude

$$\sum_{n=1}^{t-1} \tfrac{1-\gamma^{t-n}}{1-\gamma} \eta_{n+1} \nabla_{\omega_n} J(\omega_n)^T \left(\nabla_{\omega_n} J\left(\omega_n, \xi_n\right) - \nabla_{\omega_n} J\left(\omega_n, \xi_n\right)\right). \tag{S9}$$

The above formula is convergent. So B is convergent, and $g(\omega_t + 1)$ is also convergent. Substituting (S7) into (S6) we get:

$$J\left(\omega_{t+1}\right) = \zeta_t - \sum_{n=1}^{t} \eta_n \|\nabla_{\omega_n} J\left(\omega_n\right)\|^2, \tag{S10}$$

where $\{\zeta_t\}$ is defined as follows:

$$\begin{aligned}
\zeta_t = & J\left(\omega_1\right) - \tfrac{\gamma-\gamma^{t+1}}{1-\gamma} \nabla_{\omega_1} J(\omega_1)^T v_1 + \tfrac{1-\gamma^t}{1-\gamma} \eta_1 \nabla_{\omega_1} J\left(\omega_1\right) \nabla_{\omega_1} J\left(\omega_1, \xi_1\right) \\
+ & \sum_{n=1}^{t} \tfrac{1-\gamma^{t-n+1}}{1-\gamma} \eta_n \nabla_{\omega_n} J(\omega_n)^T \left(\nabla_{\omega_n} J\left(\omega_n, \xi_n\right) - \nabla_{\omega_n} J\left(\omega_n\right)\right) \\
- & \tfrac{1}{2} \sum_{n=1}^{t} v_n^T H_{\omega\omega}\left(\zeta_n\right) v_n + \sum_{n=1}^{t} \tfrac{1-\gamma^{t-n+1}}{1-\gamma} v_n^T H_{\omega\omega}\left(\zeta_n\right) v_n.
\end{aligned} \tag{S11}$$

Since $\{\zeta_t\}$ almost certainly converges, Lemma 2 shows that,

$$\omega_t \to \omega^*. \tag{S12}$$

III. Supplementary Table

Here are some supplementary tables in the Experiments section.

TABLE S1
Hyper-parameters Settings of M1-10 on D1-8

Datasets	Hyper-parameter settings				
D1	M1: Adaptive M6: λ_1=2^{-10}, λ_2=2^{-9}	M2: η=2^{-7}, λ=2^{-6} M7: λ_1=2^{-6}, λ_2=2^{-2}	M3: η=2^{-10}, λ=2^{-8} M8: λ_1=2^{-3}, λ_2=2^{-4}	M4: η=2^{-7}, λ=2^{-8} M9: η=2^{-9}, λ=2^{-10}	M5: P=1.9, λ=2^{-6} M10: α=2^{-0}, β=2^{-8}, λ=2^{-7}
D2	M1: Adaptive M6: λ_1=2^{-8}, λ_2=2^{-11}	M2: η=2^{-6}, λ=2^{-6} M7: λ_1=2^{-3}, λ_2=2^{-2}	M3: η=2^{-10}, λ=2^{-10} M8: λ_1=2^{-1}, λ_2=2^{-3}	M4: η=2^{-8}, λ=2^{-6} M9: η=2^{-8}, λ=2^{-6}	M5: P=1.9, λ=2^{-5} M10: α=2^{-0}, β=2^{-9}, λ=2^{-6}
D3	M1: Adaptive M6: λ_1=2^{-7}, λ_2=2^{-10}	M2: η=2^{-7}, λ=2^{-6} M7: λ_1=2^{-7}, λ_2=2^{-1}	M3: η=2^{-9}, λ=2^{-12} M8: λ_1=2^{-9}, λ_2=2^{-3}	M4: η=2^{-7}, λ=2^{-8} M9: η=2^{-8}, λ=2^{-2}	M5: P=1.9, λ=2^{-5} M10: α=2^{-1}, β=2^{-7}, λ=2^{-6}
D4	M1: Adaptive M6: λ_1=2^{-6}, λ_2=2^{-8}	M2: η=2^{-7}, λ=2^{-6} M7: λ_1=2^{-5}, λ_2=2^{-2}	M3: η=2^{-9}, λ=2^{-12} M8: λ_1=2^{-0}, λ_2=2^{-4}	M4: η=2^{-8}, λ=2^{-5} M9: η=2^{-7}, λ=2^{-7}	M5: P=1.9, λ=2^{-5} M10: α=2^{-0}, β=2^{-4}, λ=2^{-5}
D5	M1: Adaptive M6: λ_1=2^{-7}, λ_2=2^{-9}	M2: η=2^{-7}, λ=2^{-6} M7: λ_1=2^{-12}, λ_2=2^{-6}	M3: η=2^{-10}, λ=2^{-11} M8: λ_1=2^{-2}, λ_2=2^{-4}	M4: η=2^{-6}, λ=2^{-9} M9: η=2^{-9}, λ=2^{-11}	M5: P=1.7, λ=2^{-6} M10: α=2^{-0}, β=2^{-3}, λ=2^{-8}
D6	M1: Adaptive M6: λ_1=2^{-8}, λ_2=2^{-11}	M2: η=2^{-7}, λ=2^{-7} M7: λ_1=2^{-9}, λ_2=2^{-6}	M3: η=2^{-9}, λ=2^{-12} M8: λ_1=2^{-1}, λ_2=2^{-4}	M4: η=2^{-5}, λ=2^{-10} M9: η=2^{-9}, λ=2^{-10}	M5: P=1.9, λ=2^{-5} M10: α=2^{-2}, β=2^{-1}, λ=2^{-8}
D7	M1: Adaptive M6: λ_1=2^{-9}, λ_2=2^{-12}	M2: η=2^{-10}, λ=2^{-7} M7: λ_1=2^{-12}, λ_2=2^{-6}	M3: η=2^{-10}, λ=2^{-12} M8: λ_1=2^{-0}, λ_2=2^{-5}	M4: η=2^{-5}, λ=2^{-7} M9: η=2^{-9}, λ=2^{-9}	M5: P=1.9, λ=2^{-5} M10: α=2^{-2}, β=2^{-2}, λ=2^{-8}
D8	M1: Adaptive M6: λ_1=2^{-8}, λ_2=2^{-11}	M2: η=2^{-7}, λ=2^{-7} M7: λ_1=2^{-5}, λ_2=2^{-6}	M3: η=2^{-8}, λ=2^{-12} M8: λ_1=2^{-0}, λ_2=2^{-4}	M4: η=2^{-5}, λ=2^{-9} M9: η=2^{-8}, λ=2^{-10}	M5: P=1.9, λ=2^{-5} M10: α=2^{-2}, β=2^{-3}, λ=2^{-11}

TABLE S2
TOTAL TIME OF ALL TEST MODELS ON D1-8. (UNIT: MINUTES. NOTE THAT THE TIME INCLUDES THE TIME TO ADJUST THE HYPER-PARAMETERS.)

Models	D1	D2	D3	D4	D5	D6	D7	D8	Win/Loss
Time cost in RMSE↓									
M1	$5.8_{\pm 3.3E-01}$	$\mathbf{3.3}_{\pm 6.2E-01}$	$2.0_{\pm 2.4E-01}$	$\mathbf{0.7}_{\pm 6.1E-02}$	$28.0_{\pm 3.1E-00}$	$16.4_{\pm 7.2E-01}$	$8.1_{\pm 4.5E-01}$	$5.6_{\pm 1.0E-00}$	–
M2	$308.1_{\pm 1.3E+01}$	$71.7_{\pm 1.0E+00}$	$99.0_{\pm 2.5E+00}$	$32.0_{\pm 1.0E+00}$	$74.2_{\pm 1.1E+00}$	$74.4_{\pm 1.9E+01}$	$295.0_{\pm 2.6E+00}$	$22.3_{\pm 7.4E-01}$	8/0
M3	$306.7_{\pm 2.1E+01}$	$82.7_{\pm 3.7E-01}$	$23.7_{\pm 1.7E-01}$	$11.0_{\pm 2.1E-01}$	$28.8_{\pm 1.5E-00}$	$\mathbf{13.3}_{\pm 7.3E-02}$	$18.6_{\pm 1.7E-01}$	$1.8_{\pm 5.4E-02}$	6/2
M4	$50.5_{\pm 4.5E-00}$	$80.6_{\pm 5.6E-00}$	$22.8_{\pm 5.6E-01}$	$20.5_{\pm 3.9E-01}$	$48.9_{\pm 7.8E-01}$	$38.1_{\pm 2.1E-00}$	$5.2_{\pm 1.8E-00}$	$11.1_{\pm 5.3E-01}$	7/1
M5	$40.3_{\pm 5.2E-00}$	$43.0_{\pm 4.7E-00}$	$33.2_{\pm 3.2E-01}$	$11.5_{\pm 4.8E-01}$	$267.4_{\pm 6.3E-00}$	$125.2_{\pm 1.0E-01}$	$115.5_{\pm 5.6E-01}$	$86.1_{\pm 6.5E-00}$	8/0
M6	$\mathbf{2.2}_{\pm 6.8E-01}$	$5.9_{\pm 6.4E-01}$	$\mathbf{1.3}_{\pm 4.8E-02}$	$0.8_{\pm 7.2E-02}$	$\mathbf{3.4}_{\pm 2.7E-02}$	$17.6_{\pm 3.4E-00}$	$38.6_{\pm 2.2E-01}$	$10.5_{\pm 3.4E-01}$	5/3
M7	$67.4_{\pm 1.4E+01}$	$56.2_{\pm 6.4E-00}$	$17.4_{\pm 6.8E-00}$	$20.3_{\pm 5.5E-00}$	$139.9_{\pm 1.1E+01}$	$224.2_{\pm 6.6E-00}$	$123.9_{\pm 8.5E-00}$	$246.6_{\pm 2.1E-00}$	8/0
M8	$5.1_{\pm 1.3E-00}$	$16.6_{\pm 2.6E-00}$	$10.9_{\pm 6.0E-02}$	$1.3_{\pm 1.9E-01}$	$55.1_{\pm 1.2E+01}$	$39.0_{\pm 6.4E-00}$	$5.2_{\pm 1.3E-01}$	$\mathbf{1.6}_{\pm 1.0E-01}$	5/3
M9	$131.5_{\pm 2.8E+01}$	$45.5_{\pm 1.2E-00}$	$36.6_{\pm 9.3E-01}$	$9.7_{\pm 4.3E-01}$	$19.1_{\pm 5.1E-01}$	$16.3_{\pm 4.8E-01}$	$63.9_{\pm 1.6E-00}$	$8.8_{\pm 5.9E-01}$	6/2
M10	$222.1_{\pm 8.8E-00}$	$202.2_{\pm 5.9E-00}$	$42.8_{\pm 4.3E-00}$	$1.9_{\pm 7.6E-02}$	$127.9_{\pm 3.1E-00}$	$109.7_{\pm 5.2E-01}$	$95.2_{\pm 1.6E-00}$	$32.0_{\pm 9.7E-01}$	8/0
Time cost in MAE↓									
M1	$8.1_{\pm 2.0E-01}$	$\mathbf{4.1}_{\pm 7.8E-01}$	$2.8_{\pm 8.9E-02}$	$\mathbf{0.6}_{\pm 5.8E-02}$	$48.1_{\pm 1.2E-00}$	$26.7_{\pm 4.2E-00}$	$\mathbf{15.0}_{\pm 2.2E-00}$	$8.3_{\pm 1.0E-00}$	–
M2	$515.8_{\pm 2.2E+01}$	$133.6_{\pm 4.6E+00}$	$190.6_{\pm 5.5E+00}$	$34.7_{\pm 1.5E+00}$	$252.7_{\pm 1.4E+00}$	$192.6_{\pm 3.5E+01}$	$294.7_{\pm 3.2E+00}$	$46.3_{\pm 1.4E+00}$	8/0
M3	$328.2_{\pm 2.0E+01}$	$106.2_{\pm 1.8E-00}$	$41.9_{\pm 4.2E-01}$	$13.2_{\pm 1.9E-01}$	$117.3_{\pm 2.9E-00}$	$63.5_{\pm 1.6E-00}$	$62.7_{\pm 1.0E-00}$	$\mathbf{4.0}_{\pm 4.4E-02}$	7/1
M4	$70.2_{\pm 6.4E-00}$	$110.8_{\pm 4.6E-00}$	$34.0_{\pm 1.2E-00}$	$18.4_{\pm 4.5E-01}$	$117.4_{\pm 2.6E-00}$	$94.1_{\pm 3.1E-00}$	$22.9_{\pm 7.3E-01}$	$18.7_{\pm 6.4E-01}$	8/0
M5	$40.0_{\pm 3.7E-00}$	$35.3_{\pm 1.9E-00}$	$27.5_{\pm 8.1E-01}$	$10.9_{\pm 4.8E-01}$	$197.8_{\pm 9.0E-00}$	$91.8_{\pm 4.7E-00}$	$82.1_{\pm 5.5E-01}$	$53.2_{\pm 3.6E-00}$	8/0
M6	$\mathbf{2.2}_{\pm 6.8E-01}$	$5.3_{\pm 2.6E-02}$	$\mathbf{1.2}_{\pm 1.5E-01}$	$0.9_{\pm 7.2E-02}$	$\mathbf{3.5}_{\pm 3.2E-02}$	$17.9_{\pm 3.8E-00}$	$36.8_{\pm 7.4E-01}$	$10.6_{\pm 2.2E-01}$	4/4
M7	$153.2_{\pm 3.6E-00}$	$71.2_{\pm 2.9E-00}$	$29.7_{\pm 1.1E-00}$	$20.3_{\pm 2.4E-00}$	$145.6_{\pm 1.0E+01}$	$205.9_{\pm 1.1E+01}$	$146.9_{\pm 1.0E+01}$	$235.0_{\pm 5.5E-00}$	8/0
M8	$46.8_{\pm 6.6E-00}$	$19.2_{\pm 1.1E-00}$	$11.8_{\pm 5.0E-02}$	$7.3_{\pm 2.7E-01}$	$55.8_{\pm 7.5E-00}$	$41.6_{\pm 3.1E-00}$	$40.8_{\pm 9.0E-00}$	$12.0_{\pm 2.7E-00}$	8/0
M9	$131.0_{\pm 2.8E+01}$	$52.8_{\pm 2.6E-00}$	$50.0_{\pm 1.9E-00}$	$11.0_{\pm 3.8E-01}$	$19.1_{\pm 5.1E-01}$	$\mathbf{16.3}_{\pm 4.8E-01}$	$64.7_{\pm 1.4E-00}$	$12.9_{\pm 7.1E-01}$	6/2
M10	$389.5_{\pm 2.0E+01}$	$242.0_{\pm 2.7E-00}$	$66.7_{\pm 3.2E-00}$	$3.6_{\pm 3.0E-01}$	$127.6_{\pm 3.1E-00}$	$109.7_{\pm 5.2E-01}$	$95.2_{\pm 1.6E-00}$	$32.0_{\pm 9.7E-01}$	8/0
Time cost in R^2↓									
M1	$5.8_{\pm 3.3E-01}$	$\mathbf{3.2}_{\pm 8.2E-01}$	$2.3_{\pm 1.5E-01}$	$\mathbf{0.7}_{\pm 6.1E-02}$	$33.0_{\pm 3.1E-00}$	$16.6_{\pm 5.1E-01}$	$8.1_{\pm 3.6E-01}$	$5.6_{\pm 1.0E-00}$	–
M2	$366.2_{\pm 1.8E+01}$	$78.2_{\pm 1.1E+00}$	$123.1_{\pm 2.0E+00}$	$34.6_{\pm 1.2E+00}$	$78.7_{\pm 1.4E+00}$	$83.2_{\pm 2.3E+01}$	$294.7_{\pm 3.2E+00}$	$22.5_{\pm 8.1E-01}$	8/0
M3	$322.1_{\pm 2.3E+01}$	$101.6_{\pm 1.1E-00}$	$25.5_{\pm 2.9E-01}$	$11.6_{\pm 2.1E-01}$	$30.6_{\pm 1.3E-00}$	$\mathbf{13.9}_{\pm 3.1E-01}$	$20.3_{\pm 1.7E-01}$	$1.8_{\pm 6.6E-02}$	5/3
M4	$53.0_{\pm 3.3E-00}$	$110.7_{\pm 6.2E-00}$	$24.5_{\pm 5.6E-01}$	$20.7_{\pm 3.0E-01}$	$65.8_{\pm 7.9E-01}$	$55.0_{\pm 2.1E-00}$	$\mathbf{5.1}_{\pm 2.8E-01}$	$18.5_{\pm 6.6E-01}$	7/1
M5	$40.3_{\pm 1.2E-00}$	$42.1_{\pm 3.7E-00}$	$33.6_{\pm 6.6E-01}$	$12.0_{\pm 9.7E-01}$	$266.5_{\pm 4.7E-00}$	$124.7_{\pm 8.1E-00}$	$117.3_{\pm 3.1E-00}$	$84.4_{\pm 5.4E-00}$	8/0
M6	$\mathbf{2.2}_{\pm 6.8E-01}$	$5.9_{\pm 6.4E-01}$	$\mathbf{1.3}_{\pm 4.8E-02}$	$0.8_{\pm 7.2E-02}$	$\mathbf{3.4}_{\pm 4.4E-02}$	$17.6_{\pm 3.4E-00}$	$38.6_{\pm 1.8E-01}$	$10.4_{\pm 1.7E-01}$	5/3
M7	$72.2_{\pm 2.4E+01}$	$73.3_{\pm 4.9E-00}$	$21.0_{\pm 9.3E-00}$	$23.8_{\pm 3.6E-00}$	$139.9_{\pm 1.1E+01}$	$229.8_{\pm 8.2E-00}$	$131.8_{\pm 7.2E-00}$	$249.1_{\pm 2.8E-00}$	8/0
M8	$5.1_{\pm 1.3E-00}$	$16.6_{\pm 2.6E-00}$	$10.9_{\pm 6.0E-02}$	$1.3_{\pm 1.9E-01}$	$56.9_{\pm 1.0E+01}$	$39.0_{\pm 6.4E-00}$	$5.2_{\pm 6.4E-02}$	$\mathbf{1.6}_{\pm 1.0E-01}$	5/3
M9	$131.5_{\pm 2.8E+01}$	$46.8_{\pm 1.1E-00}$	$39.3_{\pm 3.4E-00}$	$9.9_{\pm 3.5E-01}$	$19.1_{\pm 5.1E-01}$	$16.3_{\pm 4.8E-01}$	$64.7_{\pm 1.4E-00}$	$9.3_{\pm 5.9E-01}$	6/2
M10	$233.1_{\pm 4.3E-01}$	$203.7_{\pm 5.7E-00}$	$43.5_{\pm 4.7E-00}$	$1.9_{\pm 7.6E-02}$	$127.9_{\pm 3.1E-00}$	$109.7_{\pm 5.2E-01}$	$95.2_{\pm 1.6E-00}$	$32.0_{\pm 9.7E-01}$	8/0

TABLE S3
FRIEDMAN TEST RESULTS ON ACCURACY (RMSE, MAE, R^2) AND EFFICIENCY (TIME COST IN RMSE, TIME COST IN MAE, TIME COST IN R^2)

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Accuracy	**1.00**	2.42	4.96	3.04	6.46	8.17	5.67	6.08	7.63	9.58
Efficiency	**2.25**	8.83	5.21	5.77	6.79	2.38	7.88	3.31	4.88	7.71