A Generalized Nesterov’s Accelerated Gradient-Incorporated Non-negative Latent-factorization-of-tensors Model for Efficient Representation to Dynamic QoS Data: Supplementary File

Minzhi Chen, Renfang Wang, Yan Qiao, *Senior* *Member*, *IEEE*, and Xin Luo, *Senior* *Member*, *IEEE*

This is the supplementary file for the paper entitled *Generalized Nesterov’s Accelerated Gradient-Incorporated Non-negative Latent-factorization-of-tensors Model*. Additional discussions and experimental results are put into this file.

# I. Supplementary Tables

TABLE S1. Estimation accuracy (RMSE) of M1-M6 on D1-D2, including Win/Loss and Friedman Test.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Datasets** | **M1** | **M2** | **M3** | **M4** | **M5** | **M6** |
| **D1.1** | **3.0918±1.48E-02** | 3.1557±2.25E-02 | 5.3019±3.43E-01 | 4.8965±3.51E-01 | 3.6367±3.21E-02 | 3.7408±2.31E-03 |
| **D1.2** | **2.9095±1.22E-02** | 2.9284±1.46E-02 | 4.9578±2.79E-01 | 4.8647±5.37E-02 | 3.6087±9.87E-03 | 3.6150±2.33E-03 |
| **D1.3** | **2.8245±3.40E-02** | 2.8454±1.12E-02 | 4.0871±2.92E-01 | 4.8473±1.73E-01 | 3.3477±2.95E-02 | 3.6061±3.85E-03 |
| **D1.4** | **2.8174±2.82E-02** | 2.8199±7.60E-03 | 4.0183±4.76E-01 | 4.5863±1.01E-01 | 3.2443±1.837E-02 | 3.5742±1.49E-03 |
| **D1.5** | **2.8064±2.95E-02** | 2.8070±6.50E-03 | 3.8034±4.02E-02 | 4.5133±7.22E-02 | 3.0984±1.65E-02 | 3.4947±2.32E-03 |
| **D2.1** | **28.297±7.52E-01** | 28.935±1.12E+00 | 46.818±8.29E-01 | 46.336±1.42E+00 | —— | 29.057±9.87E-01 |
| **D2.2** | **24.901±4.64E-02** | 25.059±5.16E-01 | 41.689±7.23E-01 | 45.839±1.19E+00 | —— | 27.622±8.77E-01 |
| **D2.3** | **20.637±6.73E-01** | 21.755±6.00E-01 | 38.785±5.77E-01 | 44.938±6.68E-01 | —— | 27.421±9.27E-01 |
| **D2.4** | **19.525±2.28E-01** | 20.097±1.86E-01 | 36.241±1.11E+00 | 44.246±9.23E-01 | —— | 26.060±8.19E-01 |
| **D2.5** | **19.187±1.86E-01** | 19.351±1.06E-01 | 33.653±2.53E-01 | 43.904±3.51E-01 | —— | 25.101±6.78E-01 |
| **\*Win/Loss** | —— | 10/0 | 10/0 | 10/0 | 5/0 | 10/0 |
| **Friedman Rank** | 1.000 | 2.000 | 4.800 | 5.200 | 4.500 | 3.500 |

\* Denoting that M1 (our proposed model) wins/loses when compared with the other benchmarks.

TABLE S2. Iterative count of M1-6 on D1-2, including Win/Loss and Friedman Test.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Datasets** | **M1** | **M2** | **M3** | **M4** | **M5** | **M6** |
| **D1.1** | **5±2** | 68±12 | 624±76 | 193±35 | 1000±0 | —— |
| **D1.2** | **18±3** | 385±32 | 708±33 | 338±27 | 1000±0 | —— |
| **D1.3** | **20±3** | 605±76 | 1000±0 | 420±32 | 1000±0 | —— |
| **D1.4** | **21±2** | 735±56 | 1000±0 | 510±23 | 1000±0 | —— |
| **D1.5** | **23±3** | 816±44 | 1000±0 | 552±61 | 1000±0 | —— |
| **D2.1** | **2±0** | 24±1 | 1000±0 | 173±34 | 1000±0 | —— |
| **D2.2** | **3±0** | 35±2 | 1000±0 | 249±12 | 1000±0 | —— |
| **D2.3** | **6±2** | 159±12 | 1000±0 | 917±12 | 1000±0 | —— |
| **D2.4** | **7±2** | 174±16 | 1000±0 | 1000±0 | 1000±0 | —— |
| **D2.5** | **13±2** | 316±31 | 1000±0 | 1000±0 | 1000±0 | —— |
| **\*Win/Loss** | —— | 10/0 | 10/0 | 10/0 | 10/0 | —— |
| **Friedman**  **Ranks** | 1.000 | 2.400 | 4.300 | 2.800 | 4.500 |  |

\* Denoting that M1 (our proposed model) wins/loses when compared with the other benchmarks.

TABLE S3. Time costs (Secs) of M1-6 on D1-2, including Win/Loss and Friedman Test.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Datasets** | **M1** | **M2** | **M3** | **M4** | **M5** | **M6** |
| **D1.1** | **42±17** | 51±9 | 6,003±731 | 133±24 | 721±0 | 94±3 |
| **D1.2** | **282±47** | 581±48 | 14,134±659 | 514±41 | 1,490±0 | 305±2 |
| **D1.3** | **881±132** | 2,567±322 | 50,509±0 | 1,546±118 | 4,127±0 | 932±4 |
| **D1.4** | **1,687±161** | 5,688±433 | 92,120±0 | 3,570±161 | 7,529±0 | 1,891±2 |
| **D1.5** | **3,738±488** | 12,771±689 | 186,317±0 | 7,788±861 | 14,987±0 | 3,961±5 |
| **D2.1** | 18±0 | **17**±1 | 9,414±0 | 124±24 | 689±0 | 89±3 |
| **D2.2** | **60±0** | 71±4 | 20,651±0 | 430±21 | 1,983±0 | 206±2 |
| **D2.3** | **289±96** | 787±59 | 49,426±0 | 3,658±48 | 4,793±0 | 636±5 |
| **D2.4** | **616±176** | 1,342±123 | 90,146±0 | 6,889±0 | 7,691±0 | 1,884±4 |
| **D2.5** | **1,144±356** | 2,438±488 | 182,323±0 | 13,964±0 | 15,081±0 | 2,996±6 |
| **\*Win/Loss** | —— | 9/1 | 10/0 | 10/0 | 10/0 | 10/0 |
| **Friedman Ranks** | 1.200 | 2.800 | 6.000 | 3.600 | 5.000 | 2.400 |

\* Denoting that M1 (our proposed model) wins/loses when compared with the other benchmarks .

TABLE S4. Average time cost per iteration (ms) of M1-6 on D1-2.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Datasets** | **M1** | **M2** | **M3** | **M4** | **M5** | **M6** |
| **D1.1** | 8,392 | 751 | 9,620 | 688 | 721 | —— |
| **D1.2** | 15,669 | 1,509 | 19,963 | 1,520 | 1,490 | —— |
| **D1.3** | 44,059 | 4,243 | 50,509 | 3,681 | 4,127 | —— |
| **D1.4** | 80,357 | 7,738 | 92,120 | 6,999 | 7,529 | —— |
| **D1.5** | 162,525 | 15,651 | 186,317 | 14,109 | 14,987 | —— |
| **D2.1** | 9,188 | 729 | 9,414 | 720 | 689 | —— |
| **D2.2** | 20,155 | 2,032 | 20,651 | 1,728 | 1,983 | —— |
| **D2.3** | 48,239 | 4,951 | 49,426 | 3,989 | 4,793 | —— |
| **D2.4** | 87,980 | 7,715 | 90,146 | 6,889 | 7,691 | —— |
| **D2.5** | 177,942 | 15,751 | 182,323 | 13,964 | 15,081 | —— |

TABLE S5. Results of Wilcoxon Signed-Rank Test

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Comparison** | **Accuracy** | | | **Iteration Count** | | | **Time Costs** | | |
| ***R*+** | ***R*-** | ***p*-value** | ***R*+** | ***R*-** | ***p*-value** | ***R*+** | ***R*-** | ***p*-value** |
| **M1 vs M2** | 55 | 0 | **9.77E-04** | 55 | 0 | **9.77E-04** | 54 | 1 | **2.00E-03** | |
| **M1 vs M3** | 55 | 0 | **9.77E-04** | 55 | 0 | **9.77E-04** | 55 | 0 | **9.77E-04** | |
| **M1 vs M4** | 55 | 0 | **9.77E-04** | 55 | 0 | **9.77E-04** | 55 | 0 | **9.77E-04** | |
| **M1 vs M5** | 55 | 0 | **9.77E-04** | 55 | 0 | **9.77E-04** | 55 | 0 | **9.77E-04** | |
| **M1 vs M6** | 55 | 0 | **9.77E-04** | 55 | 0 | **9.77E-04** | 47 | 8 | **2.44E-02** | |

**\***With the bold *p*-value the hypotheses that M1 is superior to other benchmarks can be accepted with a significance level of 0.05.

TABLE S6. The best values for *λ*, *λ*b, and *γ* via manual-tuned on D1-D2.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Datasets** | **Optimal hyper-parameters** | | | |
| *λ* | *λ*b | *γ* |
| **D1.1** | 2-7 | 0.5 | 0.4 |
| **D1.2** | 2-7 | 0.5 | 0.6 |
| **D1.3** | 2-7 | 0.5 | 0.6 |
| **D1.4** | 2-7 | 0.5 | 0.5 |
| **D1.5** | 2-7 | 0.5 | 0.5 |
| **D2.1** | 2-5 | 0.6 | 0.4 |
| **D2.2** | 2-5 | 0.6 | 0.4 |
| **D2.3** | 2-5 | 0.6 | 0.4 |
| **D2.4** | 2-5 | 0.6 | 0.4 |
| **D2.5** | 2-5 | 0.5 | 0.4 |

TABLE S7. Performance of GNL with/without PSO.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Datasets** | 1**Accuracy (RMSE)** | | 1**Convergence Rate (Iteration Count)** | | 1**Efficiency (#Total Time Cost in Secs)** | |
| 2Adaptive | 3Manual | Adaptive | Manual | Adaptive | Manual |
| **D1.1** | **3.0918±1.48E-02** | 3.1234±3.21E-02 | **5±2** | 135±4 | **42±17** | 26,055±772 |
| **D1.2** | **2.9095±1.22E-02** | 2.9176±2.21E-02 | **18±3** | 158±7 | **282±47** | 62,331±2,762 |
| **D1.3** | **2.8245±3.40E-02** | 2.8410±4.19E-02 | **20±3** | 169±8 | **881±132** | 183,872±8,704 |
| **D1.4** | **2.8174±2.82E-02** | 2.8156±1.39E-02 | **21±2** | 442±12 | **1,687±161** | 860,685±23,367 |
| **D1.5** | **2.8064±2.95E-02** | 2.8083±3.21E-02 | **23±3** | 762±16 | **3,738±488** | 3,023,807±63,492 |
| **D2.1** | **28.297±7.52E-01** | 28.319±6.76E-01 | **2±0** | 13±1 | 18±0 | 2,457±189 |
| **D2.2** | **24.901±4.64E-02** | 25.051±8.73E-01 | **3±0** | 15±2 | **60±0** | 9,143±1,1219 |
| **D2.3** | **20.637±6.73E-01** | 20.698±6.45E-01 | **6±2** | 46±6 | **289±96** | 58,432±7,622 |
| **D2.4** | **19.525±2.28E-01** | 20.112±8.74E-01 | **7±2** | 89±8 | **616±176** | 174,262±15,664 |
| **D2.5** | **19.187±1.86E-01** | 19.224±6.28E-01 | **13±2** | 197±12 | **1,144±356** | 787,360±47,961 |

1. The performances presented in this table include the estimation accuracy and efficiency of the tested models, where the RMSE column presents the prediction accuracy of the model, Iteration Count stands for the convergence rate of the model, and Total Time Cost records the whole training time of the model, including the cost of hyper-parameter tuning and training iteration process. For all three metrics, lower values stand for higher performance.

2. The GNL model with hyper-parameter adaptation.

3. The variant of the GNL model by removing the hyper-parameter adaptation mechanism and thus its hyper-parameters should be manually tuned. Kindly note that its performance relies on three hyper parameters, i.e., *λ*, *λb* and *γ*, which are also data-dependent. So they cannot be pre-defined, or the performance of the resultant model will be messed. The manual hyper-parameter tuning adopts the grid research, thus the total time costs of the “Manual” model include 250 (5×5×10) repetitions of the training process using different hyper-parameter combinations in the searched scale, i.e., *λ*, *λb* in the range of [2-5, 2-10] at the step of 2-1 with the exponential increase, and *γ* in the range of [0.1, 1] at the step of 0.1 with the linear increase.

# II. Supplementary Figures

D1

(a) On D1.1

D1

(b) On D1.2

D1

(c) On D1.3

D1

(d) On D1.4

D1

(e) On D1.5

Fig. S1. Training curves of M1-M6 on D1

D2

(a) On D2.1

D2

(b) On D2.2

D2

(c) On D2.3

D2

(d) On D2.4

D2

(e) On D2.5

Fig. S2. Training curves of M1-M6 on D2

RMSE_D1 RMSE_D2

(a) RMSE on D1 (b) RMSE on D2

Fig. S3. RMSE of M1-M6 on D1 and D2.

round(RMSE)_D1 round(RMSE)_D2

(a) Iteration counts on D1 (b) Iteration counts on D2

Fig. S4. Converging iteration count of M1-M6 on D1 and D2.

totaltime_D1 totaltime_D2

(a) Time costs on D1 (b) Time costs on D2

Fig. S5 Time cost of M1-M6 on D1 and D2.