Latent-Factorization-of-Tensors-Incorporated Battery Cycle Life Prediction

This is the supplementary file for the paper entitled *Latent-Factorization-of-Tensors-Incorporated Battery Cycle Life Prediction*. Additional discussions and experimental results are put into this file.

I. SUPPLEMENTARY TABLES

Algorithm I: SL_1-LFT-incorporated battery cycle life prediction framework.

Input: Known BP data consisting of L battery cells.

Operation

Initialize: LF matrixes M, N, O, P; Iteration count t=0 and *max-iteration-count* T; R, λ, and η with nonnegative constants

/*Implementing LFs representation learning by SL_1-LFT*/
while not converge **and** $t \le T$ **do**
    for each $x_{hjkl} \in \Lambda$
        for r=1 **to** R **do**
            update M according to (9)
            update N according to (9)
            update O according to (9)
            update P according to (9)
t=t+1

/* BP data completion of unknown set Γ with M, N, O, and P*/
for each $\bar{x}_{hjkl} \in \Gamma$
    for r=1 **to** R **do**
        calculate the imputed value of $\bar{x}_{hjkl}$ by (10)

/*Constructing the feature set $\boldsymbol{F}$ by merging the known and completed BP data */
/*Predicting early cycle life of BP*/
for each training instances ($F_{(l)}$, $y_{(l)}$)
    predict cycle life $\hat{y}_{(l)}$ by a data-driven cycle life prediction model

Output: Cycle life of BP

Table S1. Details of Toyota&Stanford-MIT dataset with various data densities

Data Densities	Known Counts	Unknown Counts	Battery Counts	Early Cycles	Measured Times	Measured Types
1%	173,600	17,186,400	124	100	200	7
5%	868,000	16,492,000	124	100	200	7
10%	1,736,000	15,624,000	124	100	200	7
15%	2,604,000	14,756,000	124	100	200	7
20%	3,472,000	13,888,000	124	100	200	7
30%	5,208,000	12,152,000	124	100	200	7
40%	6,944,600	10,416,000	124	100	200	7

Table S2. Accuracy of HIBP data completion on Toyota&Stanford-MIT dataset.

Data Densities	SL_1-LFT	Linear*	Zero *	Mean*	RF	KNN
1%	**38.877**$_{\pm 1.9E-06}$	47.751	53.958	51.072	51.753$_{\pm 1.8E-04}$	--
5%	**23.874**$_{\pm 4.2E-06}$	27.232	52.744	49.911	51.175$_{\pm 2.8E-04}$	51.077
10%	**17.256**$_{\pm 1.7E-06}$	22.743	51.614	48.868	50.833$_{\pm 2.6E-04}$	49.533
15%	**14.585**$_{\pm 2.5E-06}$	17.833	50.029	47.354	48.311$_{\pm 1.2E-04}$	47.882
20%	**13.452**$_{\pm 1.8E-06}$	17.499	48.520	45.925	46.512$_{\pm 1.7E-04}$	46.429
30%	**11.812**$_{\pm 1.2E-06}$	15.690	45.657	43.239	45.410$_{\pm 2.3E-04}$	43.621
40%	**10.381**$_{\pm 1.4E-06}$	12.834	42.124	39.879	29.749$_{\pm 1.5E-04}$	40.215

***Linear, Zero** and **Mean** are linear interpolation, zero and mean completion methods.

Table S3. Optimal hyper-parameters of M1 and M2

Models	Hyper-parameters' tuning ranges	Optimal values
M1	The number of trees: [6, 7, 8, 9, 10, 11, 12, 13] The number of random features: [29, 30, 31, 32]	The number of trees=6 The number of random features=29
M2	Batch size: [8, 16, 32, 64, 100] Learning rate: [0.00005, 0.00001, 0.000005, 0.000001, 0.0000005, 0.0000001]	Batch size=16 Learning rate=0.00005

Note that according to the cited references [14] and [15], the hyper-parameters tuned above are the most influential ones affecting the models' performances. The optimal hyper-parameters of each model were uniformly applied to each testing case.

Table S4. Accuracy of cycle life prediction on Toyota&Stanford-MIT dataset.

Cycle Life Prediction Models	Data Densities	Data Completion Methods					
		SL_1-LFT	Linear *	Zero *	Mean *	RF	KNN
M1	1%	**313±17**	381±23	429±16	422±15	414±19	--
	5%	**307±24**	316±25	416±23	407±22	398±20	368±21
	10%	**304±18**	311±19	388±26	350±22	379±22	357±16
	15%	**298±22**	301±21	378±20	329±23	355±17	356±18
	20%	**280±21**	290±16	334±18	310±25	324±14	326±24
	30%	**278±17**	288±21	321±14	298±18	311±12	319±5
	40%	**272±8**	281±25	305±17	287±21	283±14	302±18
M2	1%	**301±15**	315±11	322±17	315±15	321±16	--
	5%	**299±17**	312±22	318±15	314±27	319±9	315±14
	10%	**293±21**	309±24	312±9	312±23	311±6	311±13
	15%	**287±14**	302±32	309±13	308±28	309±7	310±13
	20%	**263±14**	301±19	307±12	308±12	305±24	308±21
	30%	**230±11**	299±21	303±25	306±17	303±21	306±25
	40%	**209±13**	295±17	302±21	302±13	300±12	302±11

***Linear, Zero** and **Mean** are linear interpolation, zero and mean completion, respectively.

Table S5. Details of NASA dataset with various data densities

Data Density	Known Count	Unknown Count	Battery Count	Early Cycles	Measured Times	Measured Types
1%	4,500	445,500	15	50	200	3
5%	22,500	427,000	15	50	200	3
10%	45,000	405,000	15	50	200	3
15%	67,500	382,500	15	50	200	3
20%	90,000	360,000	15	50	200	3
30%	135,000	315,000	15	50	200	3
40%	180,000	270,000	15	50	200	3

Note that NASA datasets consists of 15 battery cells with lifetimes ranging from 55 to 189 cycles. The first 50 cycles of each cell were selected to simulate its early-state degradation tendency by reasonably considering battery qualification time and accuracy. The directly measured V, I and T were selected as the degradation features. Each type of measured data is sampled 200 times on each cell. For simulating the actual HIBP data, we randomly remove a part of the measured data from the complete BP data to build a four-way HDI tensor $X^{200\times3\times50\times15}$. By reserving 1%~40% of the known BP data, seven testing cases were built to evaluate the performance of our proposed framework with various data densities.

Table S6. Accuracy of HIBP data completion on NASA data.

Data Densities	SL_1-LFT	Linear*	Zero *	Mean*	RF	KNN
1%	$\mathbf{2.388}_{\pm 2.3E-07}$	10.155	12.617	5.739	$6.000_{\pm 1.5E-05}$	--
5%	$\mathbf{1.409}_{\pm 4.6E-07}$	10.155	12.351	5.586	$5.639_{\pm 2.5E-05}$	9.956
10%	$\mathbf{1.311}_{\pm 4.6E-07}$	10.155	12.045	5.445	$5.540_{\pm 3.7E-05}$	9.380
15%	$\mathbf{1.320}_{\pm 5.2E-07}$	2.034	11.690	5.281	$5.352_{\pm 5.9E-05}$	8.967
20%	$\mathbf{1.124}_{\pm 1.7E-07}$	1.273	11.341	5.127	$5.211_{\pm 8.1E-06}$	8.654
30%	$\mathbf{0.986}_{\pm 2.1E-07}$	1.114	10.603	4.802	$4.834_{\pm 3.9E-05}$	8.041
40%	$\mathbf{0.930}_{\pm 1.8E-07}$	0.952	9.802	4.441	$4.419_{\pm 6.2E-05}$	7.414

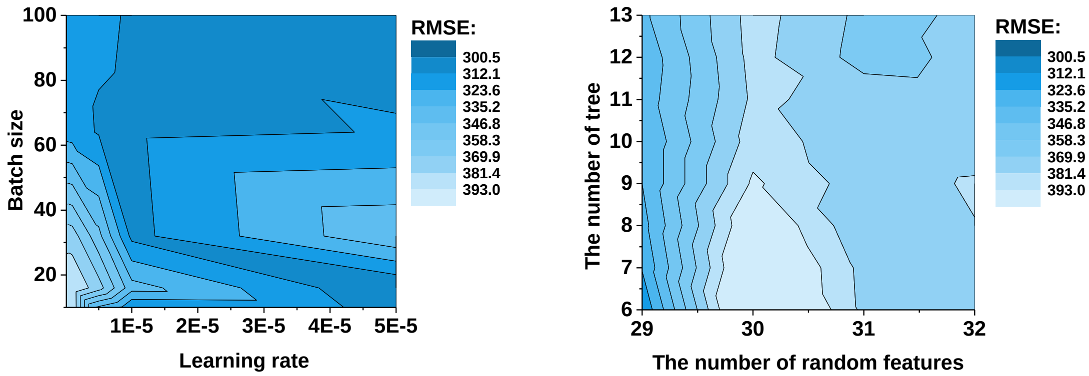
***Linear, Zero** and **Mean** are linear interpolation, zero and mean completion methods.

Table S7. Accuracy of cycle life prediction on NASA dataset.

Cycle Life Prediction Models	**Data Densities**	**Data Completion Methods**					
		SL_1-LFT	**Linear ***	**Zero ***	**Mean ***	**RF**	**KNN**
M1	1%	$\mathbf{0.712}_{\pm 5.2E-07}$	$2.685_{\pm 1.2E-07}$	$3.601_{\pm 8.1E-06}$	$3.274_{\pm 3.5E-07}$	$3.368_{\pm 6.7E-07}$	--
	5%	$\mathbf{0.490}_{\pm 1.2E-07}$	$0.859_{\pm 6.7E-07}$	$2.294_{\pm 4.5E-07}$	$2.273_{\pm 7.2E-07}$	$1.917_{\pm 5.4E-07}$	$2.250_{\pm 6.7E-07}$
	10%	$\mathbf{0.373}_{\pm 4.8E-07}$	$0.583_{\pm 4.8E-07}$	$1.619_{\pm 6.1E-07}$	$1.472_{\pm 8.2E-07}$	$1.263_{\pm 6.8E-07}$	$1.613_{\pm 5.1E-07}$
	15%	$\mathbf{0.281}_{\pm 1.4E-07}$	$0.417_{\pm 6.5E-07}$	$1.202_{\pm 3.2E-06}$	$1.042_{\pm 5.6E-07}$	$0.900_{\pm 2.5E-07}$	$0.966_{\pm 3.4E-07}$
	20%	$\mathbf{0.062}_{\pm 1.8E-07}$	$0.146_{\pm 1.7E-07}$	$0.750_{\pm 4.5E-07}$	$1.650_{\pm 5.4E-07}$	$0.339_{\pm 6.9E-07}$	$0.786_{\pm 2.1E-07}$
	30%	$\mathbf{0.044}_{\pm 2.1E-07}$	$0.259_{\pm 4.3E-07}$	$0.900_{\pm 3.1E-07}$	$0.750_{\pm 6.7E-07}$	$0.218_{\pm 6.7E-07}$	$0.375_{\pm 2.8E-07}$
	40%	$\mathbf{0.030}_{\pm 2.8E-07}$	$0.094_{\pm 8.1E-07}$	$0.120_{\pm 8.1E-07}$	$0.247_{\pm 3.1E-07}$	$0.127_{\pm 2.4E-07}$	$0.189_{\pm 6.9E-07}$
M2	1%	$\mathbf{0.683}_{\pm 4.8E-07}$	$1.102_{\pm 4.5E-07}$	$3.772_{\pm 6.4E-05}$	$4.338_{\pm 5.4E-07}$	$3.627_{\pm 6.9E-07}$	--
	5%	$\mathbf{0.382}_{\pm 5.2E-07}$	$0.988_{\pm 3.6E-07}$	$3.092_{\pm 6.3E-06}$	$2.709_{\pm 6.8E-07}$	$2.208_{\pm 5.4E-07}$	$2.348_{\pm 5.8E-07}$
	10%	$\mathbf{0.329}_{\pm 3.6E-07}$	$0.437_{\pm 2.9E-07}$	$2.055_{\pm 3.4E-07}$	$1.917_{\pm 2.1E-07}$	$1.679_{\pm 6.1E-07}$	$1.763_{\pm 2.5E-07}$
	15%	$\mathbf{0.217}_{\pm 7.1E-07}$	$0.413_{\pm 1.8E-07}$	$2.077_{\pm 8.4E-07}$	$1.498_{\pm 1.5E-07}$	$1.271_{\pm 2.4E-07}$	$1.291_{\pm 6.7E-07}$
	20%	$\mathbf{0.080}_{\pm 1.2E-07}$	$0.327_{\pm 6.7E-07}$	$1.196_{\pm 3.4E-07}$	$0.954_{\pm 1.9E-07}$	$0.924_{\pm 2.8E-07}$	$0.954_{\pm 5.2E-07}$
	30%	$\mathbf{0.045}_{\pm 1.8E-07}$	$0.286_{\pm 5.1E-07}$	$0.526_{\pm 8.0E-06}$	$0.399_{\pm 8.4E-07}$	$0.362_{\pm 5.8E-07}$	$0.417_{\pm 9.4E-07}$
	40%	$\mathbf{0.032}_{\pm 2.4E-07}$	$0.120_{\pm 6.4E-07}$	$0.307_{\pm 4.6E-05}$	$0.259_{\pm 3.1E-07}$	$0.226_{\pm 3.1E-07}$	$0.232_{\pm 1.5E-07}$

Note that 12 cells from NASA dataset were selected at random as the training set Ω, and the remaining 3 cells as the testing set Φ, and the hyper-parameters of M1 and M2 followed the optimal values as shown in Table S3. ***Linear, Zero** and **Mean** are linear interpolation, zero and mean completion, respectively.

II. Supplementary Figures



(a) M1 (b) M2

Fig. S1. The impact of hyper-parameters of M1 and M2 on Toyota&Stanford-MIT dataset with data density = 1%.

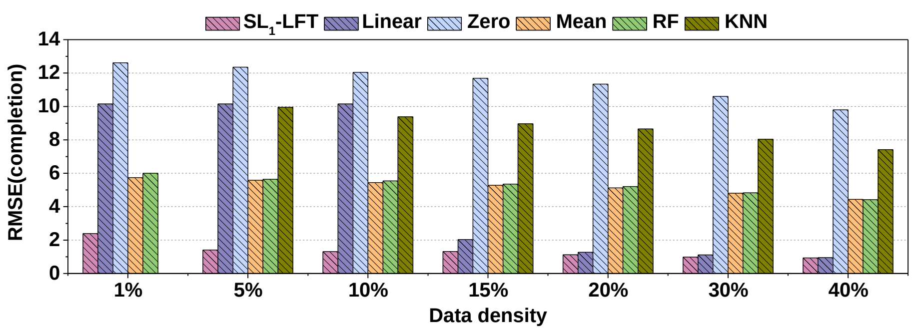
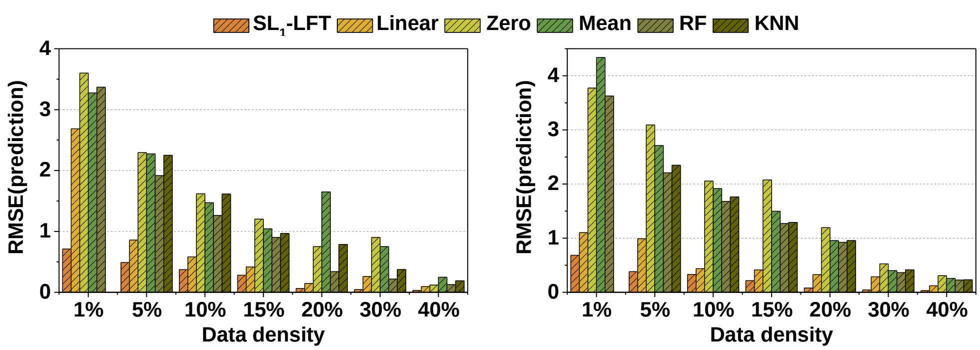


Fig. S2. Accuracy comparison of HIBP data completion on NASA dataset.



(a) M1 (b) M2

Fig. S3. Comparison of prediction accuracy of cycle life prediction models on NASA dataset.