Missing Feature Estimation-Reinforced Online Sparse Streaming Feature Selection

Anonymous

I. INTRODUCTION

This is the supplementary file for the paper entitled "*Missing Feature Estimation-Reinforced Online Sparse Streaming Feature Selection*". It mainly contains the figures of experimental results.

II. SUPPLEMENTARY TABLES

TABLE S(I) MEAN NUMBER OF SELECTED FEATURES VARYING WITH DIFFERENT ALGORITHMS, $\Psi=0.1$.

Models/Datasets	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	Average
G1+M1	7.40	12.70	14.70	18.30	23.60	8.70	36.10	44.90	43.00	78.30	54.80	11.70	29.52
G2+M1	6.50	2.00	6.00	4.60	6.00	4.00	4.00	4.70	6.00	6.40	4.10	1.00	4.61
M1	26.30	4.00	8.00	22.50	9.00	7.40	10.20	7.70	6.00	11.60	8.00	1.00	10.14
G1+M2	10.90	10.20	12.40	16.90	20.10	8.20	32.50	36.70	37.20	64.00	53.33	90.67	32.76
G2+M2	11.20	2.00	6.00	4.00	4.10	3.00	8.00	12.00	10.00	6.00	5.00	6.67	6.50
M2	75.50	3.00	31.00	74.90	16.00	18.50	16.00	37.00	24.80	54.20	61.00	47.00	38.24
G1+M3	8.20	12.70	13.80	22.50	22.20	8.67	37.00	44.30	45.60	87.20	56.67	89.67	37.38
G2+M3	197.30	29.00	22.50	15.00	27.00	1.00	1569.70	14.70	50.70	39.20	69.67	198.67	186.20
M3	102.20	16.00	1.00	1.00	1.00	27.30	125.40	1.00	1.00	1.00	1.00	1.00	23.24
G1+M4	10.90	29.10	10.70	39.00	13.70	29.80	54.20	10.60	24.20	34.70	40.00	25.33	26.85
G2+M4	11.20	27.40	9.80	20.30	15.50	33.80	49.50	10.40	30.50	39.90	37.40	34.33	26.67
M4	75.50	32.20	17.80	54.90	25.20	45.30	50.70	13.30	22.80	31.10	30.80	29.67	35.77
G1+M5	25.60	14.50	15.00	65.80	45.50	44.30	19.50	16.10	39.60	89.70	53.00	33.00	38.47
G2+M5	26.70	42.60	14.10	67.60	62.10	39.60	24.20	19.50	32.70	86.80	51.50	21.33	40.73
M5	34.40	6.50	5.20	8.80	8.20	21.40	8.10	16.20	20.90	137.30	8.90	46.70	26.88
G1+M6	28.90	7.80	4.10	6.80	5.00	10.70	11.20	3.10	3.10	5.40	9.50	14.00	9.13
G2+M6	26.10	8.00	4.60	7.30	6.00	9.90	9.00	3.00	4.00	4.40	9.40	12.33	8.67
M6	33.30	11.40	8.90	12.50	1.00	9.00	9.00	7.00	4.30	7.90	6.00	8.00	9.86

TABLE S(II) USING THE SELECTED FEATURES (RECORDED IN TABLE S(I)) TO TRAIN A CLASSIFIER FIRST AND THEN TESTING ITS ACCURACY (%), $\Psi=0.1$.

Models /Datasets	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	Average
G1+M1	$87.03_{\pm2.82}$	$82.01_{\pm3.97}$	$82.07_{\pm7.73}$	$87.06_{\pm2.13}$	$74.20_{\pm4.17}$	$73.33_{\pm4.20}$	$87.25_{\pm2.81}$	$95.05_{\pm1.56}$	$93.96_{\pm2.60}$	$97.58_{\pm0.88}$	$82.36_{\pm5.51}$	$81.10_{\pm5.16}$	$85.50_{\pm3.57}$
G2+M1	$84.87_{\pm2.63}$	$80.45_{\pm2.59}$	$78.58_{\pm2.23}$	$84.79_{\pm2.77}$	$67.56_{\pm3.28}$	$72.30_{\pm2.40}$	$86.64_{\pm1.61}$	$94.95_{\pm2.22}$	$92.87_{\pm1.71}$	$97.41_{\pm0.70}$	$81.01_{\pm3.03}$	$67.02_{\pm2.76}$	$82.37_{\pm2.33}$
M1	$71.08_{\pm2.22}$	$78.88_{\pm2.59}$	$75.09_{\pm3.84}$	$84.40_{\pm2.43}$	$56.67_{\pm2.79}$	$71.41_{\pm3.88}$	$87.05_{\pm2.46}$	$94.97_{\pm2.26}$	$92.47_{\pm2.23}$	$97.00_{\pm0.89}$	$79.29_{\pm3.24}$	$63.29_{\pm3.53}$	$79.30_{\pm2.70}$
G1+M2	$93.70_{\pm1.04}$	$82.18_{\pm4.47}$	$78.42_{\pm5.66}$	$88.06_{\pm1.89}$	$71.78_{\pm4.39}$	$73.41_{\pm5.03}$	$87.54_{\pm2.61}$	$94.76_{\pm2.47}$	$95.16_{\pm0.23}$	$97.64_{\pm0.51}$	$82.88_{\pm5.63}$	$83.14_{\pm1.11}$	$85.72_{\pm3.09}$
G2+M2	$93.62_{\pm0.91}$	$80.11_{\pm2.57}$	$76.97_{\pm4.45}$	$84.47_{\pm3.11}$	$61.25_{\pm4.13}$	$72.00_{\pm2.64}$	$86.57_{\pm1.89}$	$94.28_{\pm1.11}$	$94.53_{\pm1.61}$	$97.39_{\pm0.83}$	$81.78_{\pm5.89}$	$80.78_{\pm7.76}$	$83.65_{\pm3.08}$
M2	$88.59_{\pm0.49}$	$77.65_{\pm3.19}$	$76.51_{\pm3.32}$	$83.38_{\pm2.32}$	$61.05_{\pm5.16}$	$71.37_{\pm3.17}$	$86.23_{\pm2.80}$	$93.53_{\pm1.61}$	$93.14_{\pm2.18}$	$96.10_{\pm0.84}$	$80.15_{\pm3.83}$	$82.75_{\pm1.66}$	$82.54_{\pm2.55}$
G1+M3	$85.40_{\pm1.65}$	$80.77_{\pm3.50}$	$79.72_{\pm3.97}$	$87.76_{\pm2.34}$	$74.66_{\pm3.91}$	$72.59_{\pm4.86}$	$86.13_{\pm2.04}$	$89.57_{\pm3.58}$	$90.18_{\pm3.41}$	$97.20_{\pm0.96}$	$78.10_{\pm7.87}$	$77.25_{\pm2.72}$	$83.28_{\pm3.40}$
G2+M3	$83.61_{\pm4.78}$	$79.86_{\pm2.59}$	$63.78_{\pm18.08}$	$85.78_{\pm2.00}$	$58.57_{\pm15.01}$	$61.48_{\pm6.51}$	$83.32_{\pm2.81}$	$73.31_{\pm12.43}$	$78.79_{\pm14.10}$	$95.02_{\pm2.86}$	$77.57_{\pm14.83}$	$72.68_{\pm6.36}$	$76.15_{\pm8.53}$
M3	$80.56_{\pm3.48}$	$78.63_{\pm2.66}$	$50.13_{\pm4.49}$	$62.49_{\pm2.96}$	$45.57_{\pm1.06}$	$60.59_{\pm5.96}$	$83.26_{\pm7.68}$	$65.95_{\pm2.97}$	$62.45_{\pm5.53}$	$83.21_{\pm2.00}$	$47.53_{\pm3.83}$	$62.22_{\pm4.98}$	$65.22_{\pm3.97}$
G1+M4	$93.70_{\pm1.04}$	$78.18_{\pm5.27}$	$88.45_{\pm4.03}$	$93.07_{\pm1.29}$	$77.34_{\pm5.38}$	$68.41_{\pm3.77}$	$82.61_{\pm3.24}$	$91.54_{\pm3.17}$	$88.41_{\pm5.85}$	$96.63_{\pm1.08}$	$79.98_{\pm6.30}$	$79.22_{\pm4.14}$	$84.80_{\pm3.71}$
G2+M4	$93.62_{\pm0.91}$	$75.97_{\pm5.36}$	$85.57_{\pm5.35}$	$90.76_{\pm2.11}$	$75.74_{\pm4.28}$	$68.11_{\pm4.01}$	$81.71_{\pm3.64}$	$90.09_{\pm4.40}$	$85.87_{\pm4.79}$	$96.06_{\pm1.43}$	$78.41_{\pm5.63}$	$76.73_{\pm5.50}$	$83.22_{\pm3.95}$
M4	$88.59_{\pm0.49}$	$74.07_{\pm5.52}$	$84.50_{\pm5.74}$	$85.47_{\pm3.11}$	$71.03_{\pm5.22}$	$68.30_{\pm4.10}$	$79.81_{\pm5.46}$	$88.53_{\pm4.23}$	$83.00_{\pm5.81}$	$92.99_{\pm2.29}$	$75.89_{\pm6.80}$	$73.86_{\pm7.60}$	$80.50_{\pm4.70}$
G1+M5	$94.47_{\pm0.71}$	$84.52_{\pm4.78}$	$91.99_{\pm3.57}$	$93.10_{\pm1.28}$	$82.78_{\pm3.36}$	$81.04_{\pm3.14}$	$89.05_{\pm2.32}$	$95.96_{\pm1.96}$	$94.89_{\pm2.24}$	$97.51_{\pm0.73}$	$85.80_{\pm4.39}$	$84.05_{\pm4.22}$	$89.60_{\pm2.73}$
G2+M5	$94.19_{\pm0.80}$	$80.38_{\pm4.82}$	$89.47_{\pm4.24}$	$92.61_{\pm1.66}$	$81.75_{\pm2.92}$	$79.52_{\pm3.76}$	$87.67_{\pm2.84}$	$94.29_{\pm2.06}$	$93.11_{\pm2.34}$	$97.23_{\pm1.04}$	$85.77_{\pm4.17}$	$83.66_{\pm4.49}$	$88.30_{\pm2.93}$
M5	$81.17_{\pm1.85}$	$79.66_{\pm5.19}$	$81.58_{\pm6.76}$	$83.84_{\pm2.52}$	$69.06_{\pm5.31}$	$78.30_{\pm3.06}$	$88.94_{\pm1.81}$	$93.43_{\pm4.01}$	$88.22_{\pm3.54}$	$96.03_{\pm1.29}$	$84.41_{\pm3.31}$	$83.01_{\pm4.00}$	$83.97_{\pm3.55}$
G1+M6	$95.00_{\pm0.62}$	$83.66_{\pm4.30}$	$91.91_{\pm3.16}$	$92.99_{\pm1.48}$	$78.79_{\pm3.51}$	$70.30_{\pm4.40}$	$88.08_{\pm2.51}$	$95.44_{\pm1.87}$	$95.98_{\pm1.98}$	$97.93_{\pm0.72}$	$77.73_{\pm7.39}$	$75.03_{\pm3.65}$	$86.90_{\pm2.97}$
G2+M6	$92.23_{\pm4.39}$	$81.67_{\pm3.62}$	$88.42_{\pm4.49}$	$91.62_{\pm1.94}$	$71.99_{\pm3.02}$	$68.93_{\pm5.10}$	$86.83_{\pm1.78}$	$94.44_{\pm1.65}$	$94.18_{\pm2.00}$	$96.26_{\pm1.17}$	$75.90_{\pm6.51}$	$72.16_{\pm7.73}$	$84.55_{\pm3.62}$
M6	$81.38_{\pm3.67}$	$79.82_{\pm2.68}$	$87.10_{\pm3.47}$	$86.78_{\pm3.00}$	$49.16_{\pm1.78}$	$71.22_{\pm3.80}$	$85.32_{\pm3.61}$	$94.08_{\pm1.68}$	$93.47_{\pm2.96}$	$96.13_{\pm1.18}$	$77.37_{\pm4.52}$	$83.13_{\pm0.55}$	$82.28_{\pm2.59}$

TABLE S(III) THE RANK SUM OF THE WILCOXON SIGNED-RANKS

Models/ψ	0.3		0.5		0.7		0.9	
	*R^+	*R^-	*R^+	*R^-	*R^+	*R^-	*R^+	*R^-
G1+M1	-		-		-		-	
G2+M1	78	0	78	0	59	19	77	1
M1	78	0	78	0	71	7	67	11
G1+M2	-		-		-		-	
G2+M2	78	0	78	0	74	4	77	1
M2	78	0	78	0	71	7	60	18
G1+M3	-		-		-		-	
G2+M3	78	0	78	0	67	11	78	0
M3	78	0	78	0	78	0	78	0
G1+M4	-		-		-		-	
G2+M4	78	0	78	0	77	1	78	0
M4	76	2	78	0	77	1	69	9
G1+M5	-		-		-		-	
G2+M5	78	0	78	0	70	8	78	0
M5	78	0	78	0	78	0	67	11
G1+M6	-		-		-		-	
G2+M6	78	0	78	0	78	0	75	3
M6	70	8	67	11	62	16	69	9

* If min$\{R^+, R^-\}$>18, the null hypothesis will be taken.

TABLE S(IV) THE AVERAGE ACCURACY OF SELECTED FEATURES VARYING WITH DIFFERENT PARAMETERS OF THE MAPPING FUNCTION.

θ/Models	G1+M1	G1+M2	G1+M3	G1+M4	G1+M5	G1+M6	Average	^Rank
1.00	84.27±2.74	83.58±3.30	**83.55**±3.68	84.64±3.48	87.92±2.75	86.25±3.70	85.04±3.28	3.17
1.25	**85.49**±3.57	**85.72**±3.09	83.28±3.40	**84.80**±3.71	**89.60**±2.73	86.90±2.97	**85.97**±3.25	1.33
1.50	83.35±3.49	83.55±3.52	82.35±3.50	83.71±3.23	88.91±2.87	86.78±2.87	84.78±3.25	3.83
1.75	83.41±3.15	83.63±3.03	82.85±3.61	84.06±3.75	88.08±2.70	**87.31**±2.97	84.89±3.20	2.83
2.00	83.11±3.25	83.88±3.07	81.70±3.47	83.97±3.38	88.38±2.40	86.61±3.17	84.61±3.12	3.83

^ The Average rank.

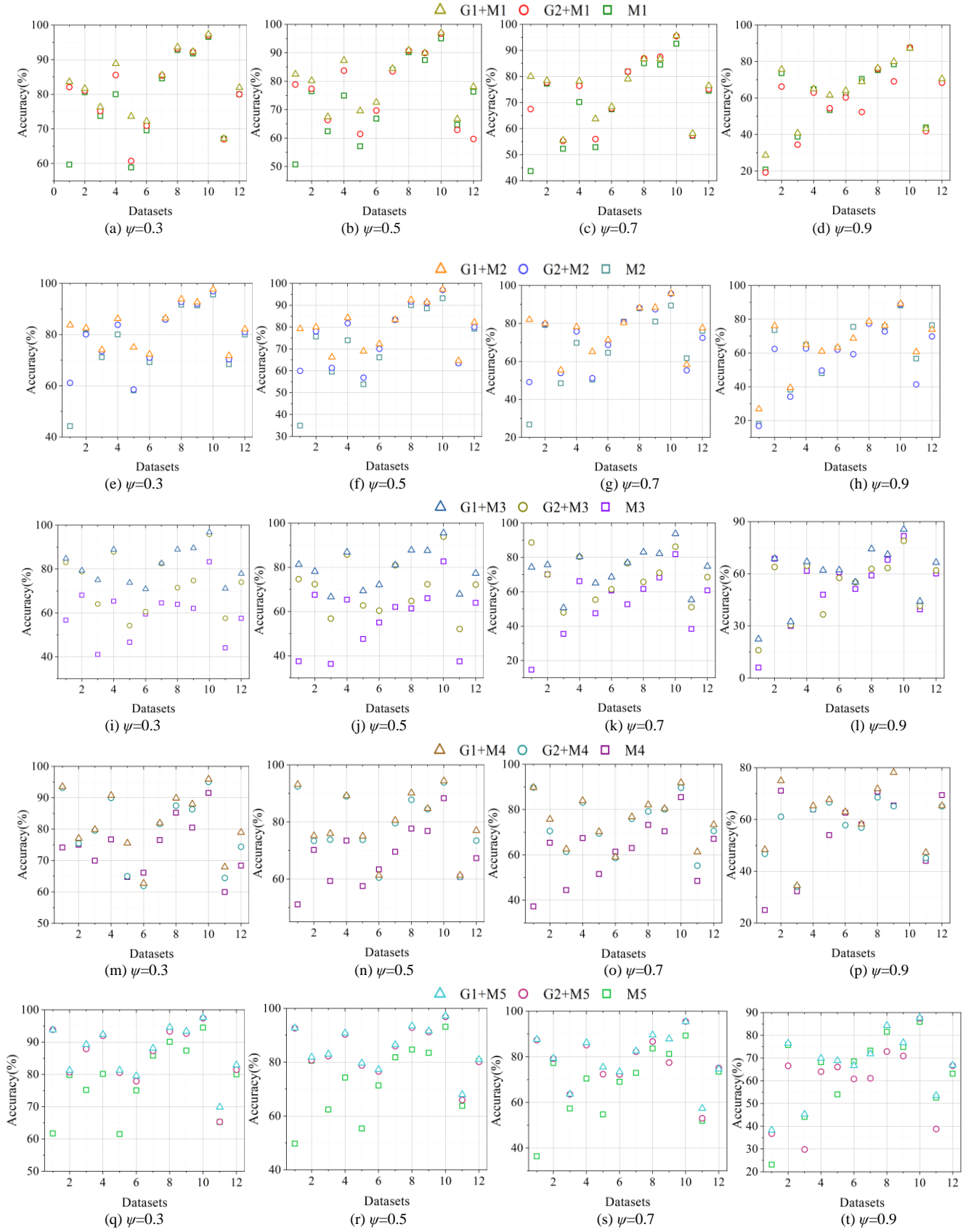TABLE S(V) THE AVERAGE ACCURACY OF SELECTED FEATURES VARYING WITH DIFFERENT CONTROLLING PARAMETERS.

η/Models	G1+M1	G1+M2	G1+M3	G1+M4	G1+M5	G1+M6	Average	^Rank
0.10	84.35±3.22	84.07±3.13	82.65±3.34	83.55±3.56	88.97±2.62	**87.15**±3.06	85.12±3.16	2.83
0.25	83.82±3.42	83.47±2.84	82.44±3.43	84.19±3.46	88.85±2.43	85.42±3.16	84.70±3.12	4.33
0.50	**85.50**±3.57	**85.72**±3.09	**83.28**±3.40	**84.80**±3.71	**89.60**±2.73	86.90±2.97	**85.97**±3.25	1.17
0.75	84.25±3.56	83.28±3.43	82.22±3.92	84.23±3.82	88.95±2.45	85.88±3.53	84.80±3.45	3.83
1.00	84.52±3.17	83.51±3.72	82.77±3.16	84.18±3.70	89.14±2.45	85.64±2.72	84.96±3.15	2.83

^ The Average rank.

TABLE S(VI) THE AVERAGE ACCURACY OF SELECTED FEATURES VARYING WITH DIFFERENT P.

P/Models	G1+M1	G1+M2	G1+M3	G1+M4	G1+M5	G1+M6	Average	^Rank
100	83.61±3.82	84.13±3.18	82.35±2.93	82.34±3.44	87.04±2.97	85.29±3.51	84.13±3.31	3.00
200	**85.50**±3.57	**85.72**±3.09	**83.28**±3.40	84.80±3.71	89.60±2.73	86.90±2.97	**85.97**±3.25	1.83
300	82.43±3.12	83.83±3.33	81.24±3.44	**85.04**±3.27	89.81±2.63	86.95±3.59	84.88±3.23	2.33
400	80.69±3.21	81.38±4.27	79.53±4.30	84.61±3.31	**89.84**±2.59	**87.74**±3.59	83.97±3.55	2.83

III. Supplementary Figures

(a) ψ=0.3   (b) ψ=0.5   (c) ψ=0.7   (d) ψ=0.9

(e) ψ=0.3   (f) ψ=0.5   (g) ψ=0.7   (h) ψ=0.9

(i) ψ=0.3   (j) ψ=0.5   (k) ψ=0.7   (l) ψ=0.9

(m) ψ=0.3   (n) ψ=0.5   (o) ψ=0.7   (p) ψ=0.9
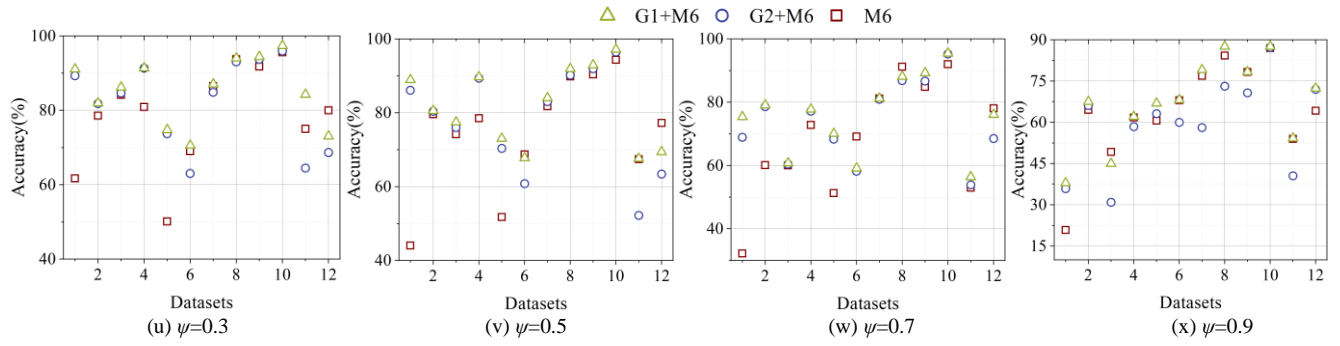
(q) ψ=0.3   (r) ψ=0.5   (s) ψ=0.7   (t) ψ=0.9

Fig. S1. The average accuracy comparison of both OS²FS and OSFS model on each dataset.
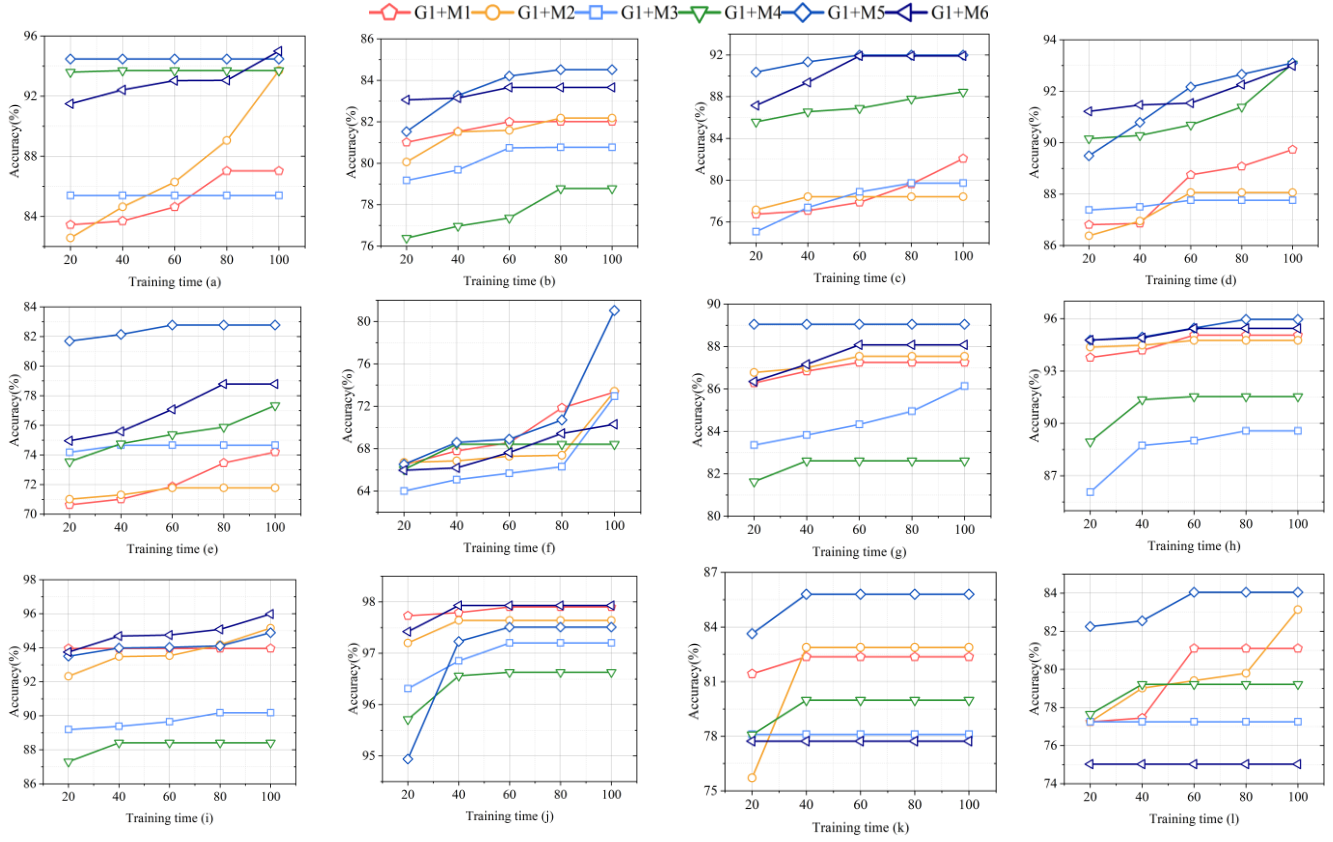
Fig. S2. Average accuracy as d increases from 20 to 100 on all the datasets at different layers. (a) D1. (b) D2. (c) D3. (d) D4. (e) D5. (f) D6. (g) D7. (h) D8. (i) D9. (j) D10. (k) D11. (l) D12.
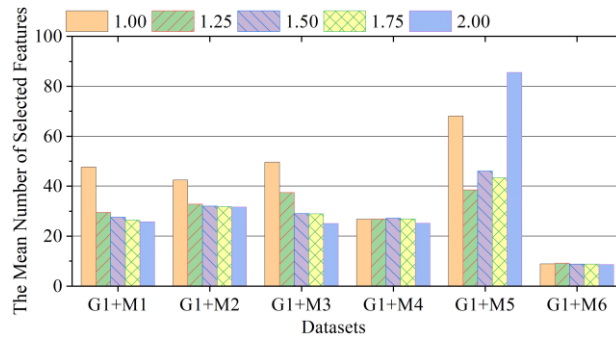

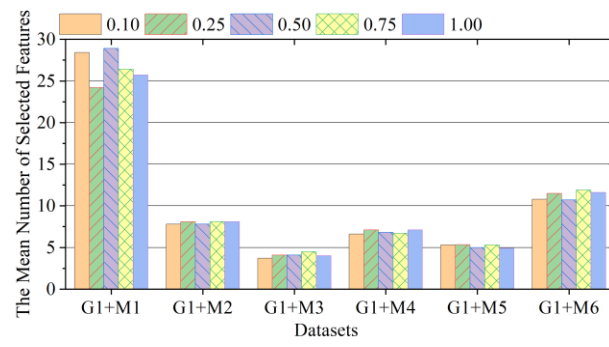Fig. S3. Mean number of selected features varying with different θ.

Fig. S4. Mean number of selected features varying with different η.