

Supplementary Material VI. The prediction accuracies obtained by three methods on $DS^{(i)}$, $DS^{(s)}$, DS_2 , evaluated by 5-fold cross-validation.

(1) The prediction accuracies obtained by the method based on chemical interactions on $DS^{(i)}$

Order	First time (%)	Second time (%)	Third time (%)	Fourth time (%)	Fifth time (%)	Mean (%)	Standard deviation (%)
1	56.37	55.95	57.31	57.47	57.92	57.00	0.82
2	21.98	24.01	22.03	22.17	22.35	22.51	0.85
3	8.91	7.25	8.90	6.90	6.84	7.76	1.06
4	5.98	5.32	4.22	5.77	5.25	5.31	0.68
5	3.16	4.19	4.11	4.41	4.56	4.09	0.55
6	2.59	2.49	2.40	2.04	1.94	2.29	0.29
7	1.47	2.38	2.51	2.49	2.51	2.27	0.45
8	1.69	1.47	1.26	1.36	1.60	1.47	0.18
9	2.37	1.13	1.48	1.36	1.25	1.52	0.49
10	0.68	1.02	1.48	1.02	0.91	1.02	0.29
11	1.24	1.25	0.80	1.24	0.91	1.09	0.22
12	1.01	1.02	1.37	1.13	1.37	1.18	0.18
13	1.35	1.25	1.03	1.24	1.14	1.20	0.12
14	0.90	0.45	0.57	0.68	0.57	0.63	0.17
15	0.56	0.57	0.91	0.79	0.68	0.70	0.15
16	0.68	0.79	0.46	0.23	0.57	0.54	0.22
17	1.13	0.79	0.68	1.24	0.46	0.86	0.32
18	0.90	0.79	0.23	1.13	0.57	0.72	0.34

19	1.13	0.57	0.91	1.02	0.68	0.86	0.23
20	0.56	1.36	1.26	0.68	1.14	1.00	0.36
21	1.01	0.68	0.68	0.79	0.80	0.79	0.14
22	1.24	1.25	1.37	1.24	1.14	1.25	0.08
23	1.13	1.02	0.68	0.90	0.91	0.93	0.16
24	0.45	0.45	0.34	0.34	0.57	0.43	0.10
25	1.01	1.36	0.57	1.02	1.03	1.00	0.28
26	0.79	1.02	1.26	1.24	1.25	1.11	0.21
27	0.56	0.68	0.34	0.79	0.46	0.57	0.18
28	0.56	0.45	0.46	0.68	0.34	0.50	0.13
29	0.34	0.91	0.46	0.57	0.68	0.59	0.22
30	0.79	0.57	0.80	0.45	0.57	0.64	0.15
31	0.00	0.11	0.11	0.00	0.23	0.09	0.10
32	0.45	0.79	0.46	0.34	0.80	0.57	0.21
33	0.68	0.45	0.46	0.57	0.57	0.54	0.09
34	0.34	0.23	0.68	0.34	0.68	0.45	0.21
35	0.45	0.57	0.57	0.45	0.34	0.48	0.10
36	0.45	0.45	0.80	0.57	1.03	0.66	0.25
37	0.68	0.79	0.34	0.68	0.34	0.57	0.21
38	0.11	0.11	0.34	0.23	0.23	0.20	0.10
39	0.45	0.45	0.46	0.57	0.68	0.52	0.10
40	0.45	0.45	0.57	0.45	0.46	0.48	0.05

41	0.23	0.34	0.34	0.34	0.34	0.32	0.05
42	0.34	0.34	0.57	0.79	0.46	0.50	0.19
43	0.34	0.23	0.68	0.11	0.34	0.34	0.21
44	0.45	0.45	0.34	0.34	0.34	0.39	0.06
45	0.45	0.79	0.80	0.68	0.68	0.68	0.14
46	0.56	0.57	0.46	0.57	0.46	0.52	0.06
47	0.23	0.23	0.34	0.11	0.34	0.25	0.10
48	0.34	0.34	0.34	0.34	0.23	0.32	0.05
49	0.11	0.11	0.00	0.00	0.00	0.05	0.06
50	0.23	0.11	0.23	0.11	0.57	0.25	0.19
51	0.90	0.45	0.68	0.79	0.68	0.70	0.17
52	0.23	0.23	0.11	0.11	0.23	0.18	0.06
53	0.34	0.34	0.34	0.45	0.34	0.36	0.05
54	0.23	0.23	0.23	0.23	0.23	0.23	0.00
55	0.00	0.00	0.23	0.00	0.00	0.05	0.10
56	0.00	0.11	0.11	0.11	0.11	0.09	0.05

(2) The prediction accuracies obtained by the method based on chemical similarities on $DS^{(s)}$

Order	First time (%)	Second time (%)	Third time (%)	Fourth time (%)	Fifth time (%)	Mean (%)	Standard deviation (%)
1	44.17	43.62	43.90	45.86	44.68	44.45	0.88
2	13.41	12.90	11.62	12.77	13.51	12.84	0.75
3	6.85	5.94	8.18	6.39	5.89	6.65	0.94
4	5.54	6.67	4.73	5.22	6.90	5.81	0.93

5	4.52	4.06	5.45	3.92	4.45	4.48	0.60
6	2.19	3.33	3.87	4.64	3.02	3.41	0.92
7	3.64	3.33	2.30	2.90	2.73	2.98	0.53
8	1.90	1.74	3.16	1.31	3.16	2.25	0.86
9	2.77	2.75	2.30	2.32	1.58	2.34	0.48
10	2.48	3.48	1.43	1.74	1.44	2.11	0.87
11	2.04	1.45	1.72	1.16	2.44	1.76	0.50
12	2.19	2.61	2.01	2.76	1.44	2.20	0.52
13	2.33	1.74	2.58	2.03	1.29	2.00	0.50
14	2.19	2.17	2.73	1.31	2.30	2.14	0.52
15	0.44	1.88	1.15	1.60	1.58	1.33	0.56
16	1.02	1.16	0.72	1.16	1.15	1.04	0.19
17	0.87	0.87	1.29	1.02	1.15	1.04	0.18
18	0.87	1.16	1.29	1.02	0.72	1.01	0.23
19	1.60	1.01	1.87	1.31	1.01	1.36	0.37
20	1.60	0.72	0.72	1.02	1.29	1.07	0.38
21	0.73	1.30	0.86	1.74	1.15	1.16	0.40
22	1.31	0.43	1.00	1.45	0.57	0.96	0.44
23	1.17	1.01	0.86	0.58	1.44	1.01	0.32
24	0.58	0.58	0.86	0.73	1.15	0.78	0.24
25	1.46	1.16	1.00	0.58	0.43	0.93	0.42
26	0.73	0.43	0.72	0.44	1.44	0.75	0.41

27	0.87	0.87	0.72	0.87	0.29	0.72	0.25
28	1.02	0.58	0.72	0.15	0.57	0.61	0.32
29	0.73	0.58	0.57	1.02	0.43	0.67	0.22
30	0.44	0.72	0.86	0.29	0.57	0.58	0.23
31	0.44	0.43	0.86	0.44	0.72	0.58	0.20
32	0.44	0.72	0.00	0.00	0.43	0.32	0.31
33	0.29	0.72	1.00	0.29	0.29	0.52	0.33
34	0.58	0.29	1.00	0.29	0.29	0.49	0.31
35	0.44	0.29	0.14	0.29	0.29	0.29	0.10
36	0.29	0.29	0.43	1.16	0.57	0.55	0.36
37	0.87	0.43	0.72	0.73	0.43	0.64	0.20
38	0.29	0.29	0.43	1.16	0.43	0.52	0.36
39	0.58	0.72	0.57	0.58	0.43	0.58	0.10
40	0.44	1.16	1.15	1.02	0.72	0.90	0.31
41	0.44	0.29	0.14	0.44	0.86	0.43	0.27
42	0.29	0.43	0.14	0.29	0.57	0.35	0.16
43	0.73	0.58	0.14	0.58	1.15	0.64	0.36
44	0.44	0.43	0.43	0.58	0.29	0.43	0.10
45	0.29	0.14	0.14	0.29	0.14	0.20	0.08
46	0.29	0.43	0.29	0.29	0.57	0.38	0.13
47	0.29	0.43	0.14	0.87	0.43	0.43	0.27
48	0.29	0.43	0.43	0.29	0.14	0.32	0.12

49	0.29	0.29	0.14	0.29	0.29	0.26	0.07
50	0.00	0.58	0.57	0.15	0.57	0.37	0.28
51	0.44	0.43	0.29	0.29	0.29	0.35	0.08
52	0.44	0.43	0.29	0.29	0.72	0.43	0.18
53	0.29	0.14	0.29	0.58	0.72	0.40	0.24
54	0.73	0.58	0.57	0.73	0.43	0.61	0.12
55	0.44	0.72	0.57	0.73	0.43	0.58	0.15
56	0.15	0.00	0.29	0.15	0.14	0.14	0.10

(3) The prediction accuracies obtained by the integrated method on DS_2

Order	First time (%)	Second time (%)	Third time (%)	Fourth time (%)	Fifth time (%)	Mean (%)	Standard deviation (%)
1	51.05	50.54	51.37	52.38	52.07	51.48	0.75
2	18.25	19.14	17.42	18.05	18.44	18.26	0.62
3	8.01	6.68	8.58	6.68	6.42	7.27	0.96
4	5.79	5.91	4.45	5.53	5.98	5.53	0.63
5	3.75	4.13	4.70	4.20	4.51	4.26	0.37
6	2.42	2.86	3.05	3.18	2.42	2.78	0.36
7	2.42	2.80	2.42	2.67	2.61	2.58	0.17
8	1.78	1.59	2.10	1.34	2.29	1.82	0.38
9	2.54	1.84	1.84	1.78	1.40	1.88	0.41
10	1.46	2.10	1.46	1.34	1.14	1.50	0.36
11	1.59	1.34	1.21	1.21	1.59	1.39	0.19
12	1.53	1.72	1.65	1.84	1.40	1.63	0.17

13	1.78	1.46	1.72	1.59	1.21	1.55	0.23
14	1.46	1.21	1.53	0.95	1.34	1.30	0.23
15	0.51	1.14	1.02	1.14	1.08	0.98	0.27
16	0.83	0.95	0.57	0.64	0.83	0.76	0.16
17	1.02	0.83	0.95	1.14	0.76	0.94	0.15
18	0.89	0.95	0.70	1.08	0.64	0.85	0.18
19	1.34	0.76	1.34	1.14	0.83	1.08	0.27
20	1.02	1.08	1.02	0.83	1.21	1.03	0.14
21	0.89	0.95	0.76	1.21	0.95	0.95	0.16
22	1.27	0.89	1.21	1.34	0.89	1.12	0.21
23	1.14	1.02	0.76	0.76	1.14	0.97	0.19
24	0.51	0.51	0.57	0.51	0.83	0.58	0.14
25	1.21	1.27	0.76	0.83	0.76	0.97	0.25
26	0.76	0.76	1.02	0.89	1.34	0.95	0.24
27	0.70	0.76	0.51	0.83	0.38	0.64	0.19
28	0.76	0.51	0.57	0.45	0.45	0.55	0.13
29	0.51	0.76	0.51	0.76	0.57	0.62	0.13
30	0.64	0.64	0.83	0.38	0.57	0.61	0.16
31	0.19	0.25	0.45	0.19	0.45	0.31	0.13
32	0.45	0.76	0.25	0.19	0.64	0.46	0.24
33	0.51	0.57	0.70	0.45	0.45	0.53	0.11
34	0.45	0.25	0.83	0.32	0.51	0.47	0.22

35	0.45	0.45	0.38	0.38	0.32	0.39	0.05
36	0.38	0.38	0.64	0.83	0.83	0.61	0.22
37	0.76	0.64	0.51	0.70	0.38	0.60	0.15
38	0.19	0.19	0.38	0.64	0.32	0.34	0.18
39	0.51	0.57	0.51	0.57	0.57	0.55	0.03
40	0.45	0.76	0.83	0.70	0.57	0.66	0.15
41	0.32	0.32	0.25	0.38	0.57	0.37	0.12
42	0.32	0.38	0.38	0.57	0.51	0.43	0.10
43	0.51	0.38	0.45	0.32	0.70	0.47	0.15
44	0.45	0.45	0.38	0.45	0.32	0.41	0.06
45	0.38	0.51	0.51	0.51	0.45	0.47	0.06
46	0.45	0.51	0.38	0.45	0.51	0.46	0.05
47	0.25	0.32	0.25	0.45	0.38	0.33	0.08
48	0.32	0.38	0.38	0.32	0.19	0.32	0.08
49	0.19	0.19	0.06	0.13	0.13	0.14	0.05
50	0.13	0.32	0.38	0.13	0.57	0.31	0.19
51	0.70	0.45	0.51	0.57	0.51	0.55	0.10
52	0.32	0.32	0.19	0.19	0.45	0.29	0.11
53	0.32	0.25	0.32	0.51	0.51	0.38	0.12
54	0.45	0.38	0.38	0.45	0.32	0.39	0.05
55	0.19	0.32	0.38	0.32	0.19	0.28	0.09
56	0.06	0.06	0.19	0.13	0.13	0.11	0.05