Supporting Information

Modified particle swarm optimization algorithm for adaptively configuring globally optimal classification and regression trees

Yan-Ping Zhou[1,2*], Li-Juan Tang[2], Jian Jiao[1], Dan-Dan Song[1], Jian-Hui Jiang[2*], Ru-Qin Yu[2*]

1. Key Laboratory of Pesticide and Chemical Biology of Ministry of Education, College of Chemistry, Central China Normal University, Wuhan 430079, P. R. China

2. State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, Hunan University, Changsha 410082, P. R. China

Table 1 Node information associated with CART configured by recursive partition for flavonoid derivatives

node	splitting variable	splitting value	node size	node output
1	S_dO	12.3917	80	3.25
2	Dipole Y	3.9915	42	3.71
3	Jurs-WNSA-1	122.211	38	2.73
4	Quadrupole xy	0.3115	9	3.00
5	S_dssC	0.004447	33	3.91
6	-	-	2	3.23
7	Jurs-WNSA-1	124.444	36	2.70
8	Jurs-WPSA-1	169.565	6	2.75
9	-	-	3	3.48
10	Jurs-RPCS	1.32529	7	4.48
11	Quadrupole yy	-20.5065	26	3.75
12	-	-	1	2.80
13	Octupole xxz		35	2.70
14	-	-	1	3.01
15	-	-	5	2.70
16	-	-	3	4.01
17	-	-	4	4.83
18	He	-17.96	12	4.01
19	Octupole xzz	-15.655	14	3.54
20	-	-	34	2.70
21	-	-	1	2.73
22	-	-	3	4.59
23	Jurs-TASA	372.83	9	3.81
24	-	-	2	2.85
25	RadiusOfGyration	4.61492	12	3.65
26	Octupole xzz	5.147	8	3.70
27	-	-	1	4.57
28	Jurs-FNSA-3	-0.15117	9	3.54
29	-	-	3	4.01
30	Jurs-FNSA-3	-0.102	5	3.87
31	-	-	3	3.45
32	-	-	2	3.86
33	Shadow ratio	4.37619	7	3.45
34	-	-	4	3.92
35	-	-	1	3.68
36	Jurs-FNSA-3	-0.11403	5	3.50
37	-	-	2	3.21
38	-	-	4	3.50
39	-	-	1	3.70

Symbol “-” represents the leaf nodes without splitting variables and values.

Table 2 Node information associated with CART configured by recursive partition for EGFR tyrosine kinase inhibitors

node	splitting variable	splitting value	node size	node output
1	Fh2o	-21.0855	49	3.91
2	Dipole_Y	-2.667	32	8.52
3	Fh2o	-7.2165	17	6.77
4	Dipole_Z	5.2825	7	7.30
5	Foct	-26.5935	25	8.86
6	Shadow_XYfrac	0.6725	13	7.10
7	-	-	4	5.70
8	-	-	4	6.46
9	-	-	3	8.41
10	Dipole_X	1.8725	16	9.36
11	RadOfGyRation	3.9455	9	7.96
12	LOMO	1.6125	11	7.31
13	-	-	2	5.97
14	-	-	2	7.54
15	Shadow_nu	2.0795	14	9.63
16	-	-	2	7.26
17	RadOfGyRation	4.203	7	8.16
18	Dipole_mag	8.8945	9	7.45
19	-	-	2	6.69
20	Dipole_Y	-1.674	7	9.14
21	S_sCH3	3.2056	7	10.11
22	Shadow_nu	2.07	5	8.34
23	-	-	2	7.73
24	Dipole_Z	2.7315	6	7.32
25	-	-	3	7.70
26	-	-	1	8.42
27	Shadow_YZfrac	0.642	6	9.26
28	S_aasC	3.538	5	9.79
29	-	-	2	10.91
30	-	-	2	8.08
31	-	-	3	8.51
32	-	-	4	7.17
33	-	-	2	7.61
34	-	-	2	9.54
35	-	-	4	9.12
36	-	-	3	9.63
37	-	-	2	10.03

Symbol “-” represents the leaf nodes without splitting variables and values.