Replication of Tool Building Project：Configuration Performance Learning

By En Li

1. Introduction

Linear regression model is easy to understand and implement and has low computational cost. However, many of the relationships between configuration parameters and performance in the data set of Lab2 are nonlinear, and linear regression cannot handle such nonlinear relationships. In addition, linear regression is easily affected by outliers and multicollinearity, and overfit or underfit. This report discusses how to optimize the linear regression algorithm and design a solution to improve the defects of the linear regression.

2. Datasets and python

(1) Datasets: Using the datasets in lab2

(2) Python: Python 3.12.9 and higher

3. Executing scripts to get results

(1) Solution results

Executing lab2_coursework.py without parameters will get the results as 5.2 and 5.3 in report, which will be saved in file lab2.csv.

(2) Experiment of HuberRegressor only

Executing lab2_huber.py without parameters will get the results in section 5.4 of report, which will be saved in file lab2_huber.csv.

4. Solution results

Table 1. Experiment results

SN	System	Dataset	Baseline (Linear Regression)			Solution		
			MAPE	MAE	RMSE	MAPE	MAE	RMSE
1	batlik	corona.csv	0.12	0.18	0.26	0.03	0.05	0.07
2	batlik	cranium.csv	0.03	0.07	0.1	0.02	0.04	0.05
3	batlik	cubus.csv	0.04	0.24	0.34	0.02	0.11	0.14
4	batlik	flame.csv	0.03	0.06	0.08	0.02	0.03	0.04
5	batlik	france.csv	0.04	0.15	0.2	0.02	0.08	0.1
6	batlik	frisco.csv	0.05	0.25	0.34	0.03	0.13	0.18
7	batlik	jelly.csv	0.17	0.32	0.44	0.03	0.05	0.08
8	batlik	karte.csv	0.15	1.49	1.9	0.03	0.33	0.48
9	batlik	mandelbrot.csv	0.05	0.06	0.09	0.02	0.03	0.03
10	batlik	strawberry.csv	0.05	0.09	0.13	0.02	0.03	0.04
11	batlik	village.csv	0.57	1.79	2.44	0.02	0.11	0.26
12	dconvert	jpeg-large.csv	0.13	0.42	0.57	0.04	0.15	0.26
13	dconvert	jpeg-medium.csv	0.1	0.18	0.24	0.04	0.06	0.11
14	dconvert	jpeg-small.csv	0.04	0.02	0.02	0.03	0.01	0.02
15	dconvert	png-large.csv	0.09	0.22	0.28	0.03	0.08	0.11
16	dconvert	png-medium.csv	0.04	0.02	0.03	0.03	0.02	0.02
17	dconvert	png-small.csv	0.03	0.01	0.02	0.03	0.01	0.02
18	dconvert	psd-large.csv	0.01	0.53	0.85	0.01	0.37	0.88
19	dconvert	psd-medium.csv	0.05	0.07	0.09	0.02	0.03	0.04
20	dconvert	psd-small.csv	0.04	0.02	0.02	0.03	0.02	0.02

21	dconvert	svg-large.csv	0.02	0.03	0.04	0.02	0.03	0.04
22	dconvert	svg-medium.csv	0.02	0.03	0.03	0.02	0.03	0.03
23	dconvert	svg-small.csv	0.04	0.02	0.03	0.03	0.01	0.02
24	h2	smallbank-1.csv	0.06	1117.19	1342.34	0.04	736.59	950.24
25	h2	smallbank-10.csv	0.05	633.98	764.34	0.02	311.96	412.29
26	h2	tpcc-2.csv	0.04	30.38	37.51	0.03	20.08	25.83
27	h2	tpcc-8.csv	0.2	109.56	165.1	0.23	92.41	194.91
28	h2	voter-16.csv	0.08	2429.21	2708.25	0.02	669.92	899.83
29	h2	voter-2.csv	0.08	2523	2803.66	0.02	643.9	891.64
30	h2	ycsb-2400.csv	0.09	1261.9	1548.1	0.05	667.97	864.99
31	h2	ycsb-600.csv	0.08	1686.9	2096.59	0.04	845.2	1135.45
32	jump3r	beethoven.wav.csv	0.25	2.24	2.89	0.19	1.83	2.56
33	jump3r	dual-channel.wav.csv	0.23	0.6	0.83	0.17	0.52	0.76
34	jump3r	helix.wav.csv	0.24	1.06	1.45	0.18	0.9	1.32
35	jump3r	single-channel.wav.csv	0.19	0.32	0.42	0.15	0.28	0.39
36	jump3r	speech.wav.csv	0.2	0.8	1.01	0.15	0.65	0.9
37	jump3r	sweep.wav.csv	0.12	0.07	0.09	0.1	0.06	0.08
38	kanzi	ambivert.csv	1.4	16.97	27.66	0.71	13.94	28.94
39	kanzi	article.csv	0.46	0.2	0.29	0.24	0.15	0.3
40	kanzi	deepfield.csv	17.67	18.79	29.05	4.83	13.76	29.57
41	kanzi	enwik8.csv	2.59	39.53	62.36	0.96	29.69	66.54
42	kanzi	fannie_mae_500k.csv	4.06	34.79	53.9	1.17	25.09	58.98
43	kanzi	large.csv	1.39	4.38	6.89	0.49	3.24	7.62
44	kanzi	misc.csv	0.46	0.4	0.61	0.2	0.29	0.66
45	kanzi	silesia.csv	2.87	82.26	129.55	0.92	61.53	142.62
46	Kanzi	vmlinux.csv	28.81	162.76	220.94	9.59	93.01	191.24
47	lrzip	ambivert.wav.tar.csv	1.49	15.16	25.55	0.61	11.6	25.73
48	lrzip	artificial.tar.csv	14.09	1.28	1.44	13.82	1.41	1.78
49	lrzip	deepfield.tar.csv	1.47	14.66	24.62	0.64	11.44	24.93
50	lrzip	enwik8.tar.csv	1.04	19.54	29.14	0.55	15.6	29.83
51	lrzip	fannie_mae_500k.tar.csv	0.12	5.46	8.14	0.1	4.89	8.27
52	lrzip	large.tar.csv	0.98	2.35	3.11	0.9	2.41	3.32
53	lrzip	misc.tar.csv	2.73	1.28	1.47	2.61	1.38	1.75
54	lrzip	silesia.tar.csv	1.18	33.22	50.49	0.61	26.94	52.02
55	lrzip	uiq-32.bin.csv	1.91	56.16	85.19	0.9	45.4	89.1
56	lrzip	uiq2-16.bin.csv	1.73	30.54	46.2	0.82	24.11	47.53
57	lrzip	uiq2-4.bin.csv	1.19	7.52	11.92	0.69	6.21	11.98
58	lrzip	uiq2-8.bin.csv	1.42	15.27	23.35	0.74	12.21	23.69
59	lrzip	vmlinux-5.10.tar.csv	0.78	4.85	7.38	0.54	4.24	7.44
60	x264	blue_sky_1080p25_short.y4m.csv	0.79	3.13	4.06	0.23	1.39	3.24
61	x264	Johnny_1280x720_60 _short.y4m.csv	0.9	2.57	3.35	0.26	1.16	2.66
62	x264	Netflix_Crosswalk_4096x2160_	0.79	8.7	11.26	0.23	3.84	8.9

		60fps_10bit_420_short.y4m.csv						
63	x264	pedestrian_area_ 1080p25_short.y4m.csv	0.72	3.56	4.61	0.2	1.56	3.79
64	x264	riverbed_1080p25_short.y4m.csv	0.66	6.64	8.78	0.2	3.05	7.12
65	x264	sd_bridge_close_cif_ short.y4m.csv	0.52	0.14	0.18	0.16	0.06	0.13
66	x264	sd_city_4cif_short.y4m.csv	0.83	1.48	1.9	0.21	0.61	1.56
67	x264	sd_crew_cif_short.y4m.csv	0.62	0.24	0.32	0.18	0.11	0.25
68	x264	sintel_trailer_2k_720p24 _short.y4m.csv	0.52	0.37	0.47	0.2	0.19	0.36
69	xz	ambivert.wav.tar.csv	1.4	13.08	17.11	0.3	2.63	8.91
70	xz	artificial.tar.csv	0.29	0.01	0.01	0.33	0.01	0.01
71	xz	deepfield.tar.csv	0.62	5.59	7.52	0.15	1.54	4.57
72	xz	enwik8.tar.csv	2.11	71.29	93.05	0.63	22.11	60.07
73	xz	fannie_mae_500k.tar.csv	0.95	29.96	35.74	0.32	10.05	21.84
74	xz	large.tar.csv	0.44	1.82	2.24	0.18	0.65	1.17
75	xz	misc.tar.csv	0.23	0.09	0.12	0.1	0.04	0.08
76	xz	silesia.tar.csv	0.99	61.95	78.1	0.28	16.45	43.26
77	xz	uiq-32.bin.csv	1.25	71.07	90.1	0.34	10:15 p.m.	50.05
78	xz	uiq2-16.bin.csv	1.1	31.8	42.62	0.35	11.36	25.84
79	xz	uiq2-4.bin.csv	0.62	4.2	5.69	0.23	1.64	3.55
80	xz	uiq2-8.bin.csv	0.84	11.59	15.33	0.31	4.67	9.3
81	xz	vmlinux-5.10.tar.csv	0.28	1.83	2.23	0.09	0.59	1.11
82	z3	AUFNIRA_z3.637557.smt2.csv	7.65	88.11	99.32	2.03	27.39	31.01
83	z3	LRA_formula_277.smt2.csv	0.11	0.32	0.32	0	0.01	0.02
84	z3	QF_AUFBV_891_sqlite3.smt2.csv	5.38	10.21	13.86	2.41	5.24	9.9
85	z3	QF_AUFLIA_swap_t1_pp_ nf_ai_00010_002.cvc.smt2.csv	0.13	0.92	1.13	0.09	0.58	0.75
86	z3	QF_BV_bench_3176.smt2.csv	10.88	133.15	152.1	3.44	40.39	50.6
87	z3	QF_BV_bench_935.smt2.csv	18.8	93.15	96.85	1.28	11.55	20.85
88	z3	QF_LIA_tightrhombus.csv	0.7	0.59	0.8	0.03	0.06	0.14
89	z3	QF_LRA_clocksynchro_7clocks. worst_case_skew.induct.smt2.csv	0.06	0.74	0.87	0.02	0.28	0.3
90	z3	QF_NRA_hong_9.smt2.csv	0	0.01	0.02	0	0.01	0.02
91	z3	QF_RDL_orb08_888.smt2.csv	0.03	0.27	0.33	0.01	0.11	0.14
92	z3	QF_UFLIA_xs_23_43.smt2.csv	0.38	8.14	9.77	0.12	2.95	3.94
93	z3	QF_UF_PEQ018_size6.smt2.csv	10.1	158.84	198.75	7.21	90.37	96.94

(1) Wilcoxon test results

The results of the Wilcoxon test in Table 2 show that all p-values are much less than 0.05, which means that the solution results significantly exceed those of the baseline.

Table 2. Wilcoxon test results

Metrics	W statistic	p-value	Explanation of the results

	(Solution vs. Baseline)	(Solution vs. Baseline)	
MAPE	40.00	1.44924E-15	Solution results are significantly better than baseline's
MAE	48.00	4.23437E-15	Solution results are significantly better than baseline's
RMSE	897.50	2.79251E-05	Solution results are significantly better than baseline's

(2) The count of better results

The data in Table 3 indicates that among the 93 sets of MAPE, MAE, and RMSE result data of the solution, 86, 83, and 63 results are better than the baseline results, respectively. In other words, among results, 92% of MAPE, 89%of MAE, and 68% of RMSE solution results are better than the baseline results.

Table 3. The number of better results (solution < Baseline)

Metrics	Count of total results	Count of better results (solution < Baseline)	Win Ratio (%)
MAPE	93	86	92%
MAE	93	83	89%
RMSE	93	63	68%

(3) The average value of the results

The data in Table 4 indicates that average MAPE, average MAE, average RMSE of the solution results are 40%, 42%, and 50% of the baseline results respectively. The solution results are significantly smaller than the baseline results, which means that the solution results are more accurate than the baseline.

Table 4. Average MAPE, Average MAE, Average RMSE

Metrics	Baseline results	Solution results	Win Ratio (%) (Solution results / Baseline results)
Average MAPE	1.78	0.72	40%
Average MAE	120.39	50.40	42%
Average RMSE	143.33	72.16	50%

5. Conclusion

According to the experimental results, the solution model significantly outperforms the baseline linear regression. Among results, 92% of MAPE, 89%of MAE, and 68% of RMSE solution results are better than the baseline results, and the average MAPE, average MAPE, average RMSE of solution results 40%, 42%, and 50% of the baseline results, indicating that the prediction error of solution model is smaller.