Machine Learning techniques to Model Data Intensive Application Performance

A. Battistello P. Ferretti

Contents

1 Fixed Datasize, $ncores^{-1}$, all the features

1.1 Query R1

1.1.1 Query R1, Datasize 250GB — testing on 40, 60 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0822	0.8237	5400	0.0577
Linear SVR	0.0727	0.9154	6305	0.0756
Polynomial SVR (2)	0.8276	0.0265	67666	0.7742
Polynomial SVR (3)	0.2752	0.0647	22513	0.2176
Polynomial SVR (4)	0.4257	0.0788	42777	0.4557
Polynomial SVR (6)	1.2206	0.0121	61724	0.7131
Gaussian SVR	0.2493	0.4372	15823	0.1991

Table 1: Results for R1 (250GB), testing on 40, 60 cores

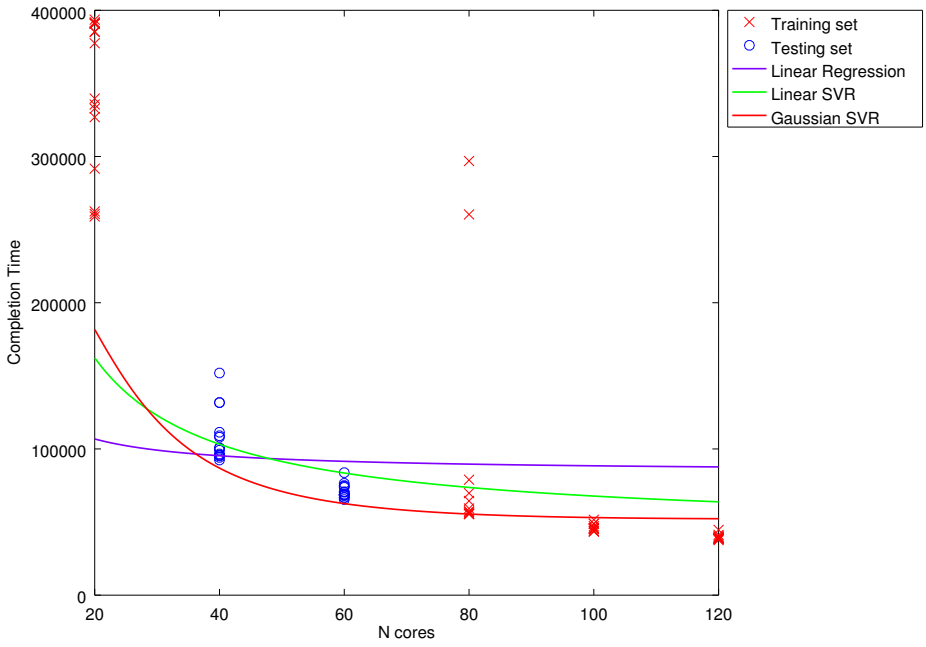

Figure 1: Completion time vs ncores for R1 (250GB), testing on 40, 60 cores

1.1.2 Query R1, Datasize 250GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2064	0.7968	9404	0.0786
Linear SVR	0.1082	0.9917	8377	0.0998
Polynomial SVR (2)	1.3488	0.1066	63412	0.6394
Polynomial SVR (3)	3.8306	0.0556	91210	1.1189
Polynomial SVR (4)	5.4272	0.0091	139814	1.6235
Polynomial SVR (6)	17.3562	0.0002	362503	4.7635
Gaussian SVR	0.2824	0.6710	18395	0.2102

Table 2: Results for R1 (250GB), testing on 60, 80 cores

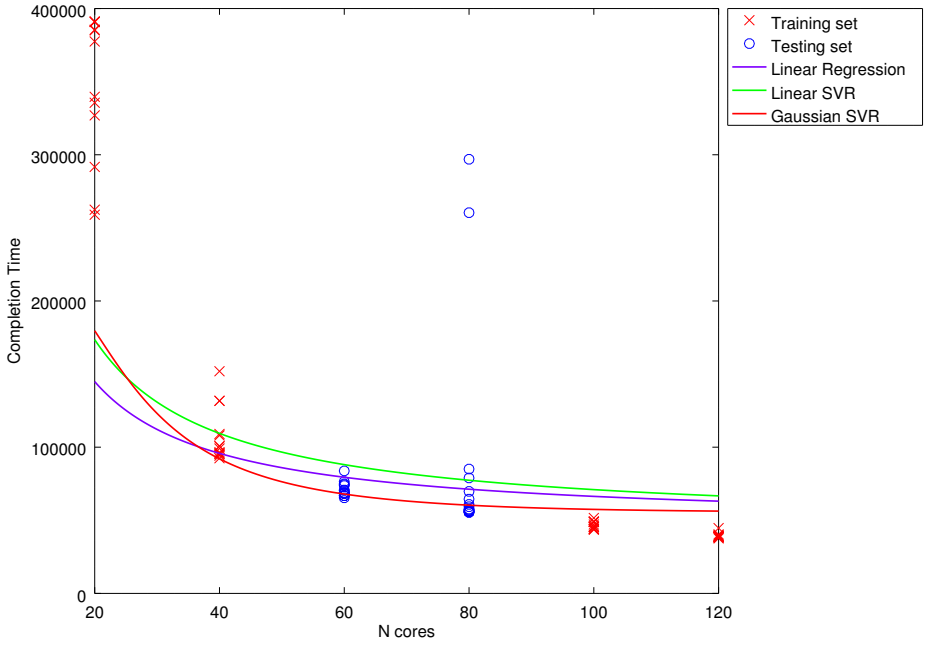

Figure 2: Completion time vs ncores for R1 (250GB), testing on 60, 80 cores

1.1.3 Query R1, Datasize 250GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1585	0.8912	10652	0.1441
Linear SVR	0.1191	0.9903	12052	0.1989
Polynomial SVR (2)	1.2612	0.8615	53979	0.5587
Polynomial SVR (3)	0.1766	0.8698	13289	0.2021
Polynomial SVR (4)	0.8678	0.6872	44057	0.5354
Polynomial SVR (6)	0.5367	0.3985	38351	0.5643
Gaussian SVR	0.2557	0.9140	15890	0.2239

Table 3: Results for R1 (250GB), testing on 80, 100 cores

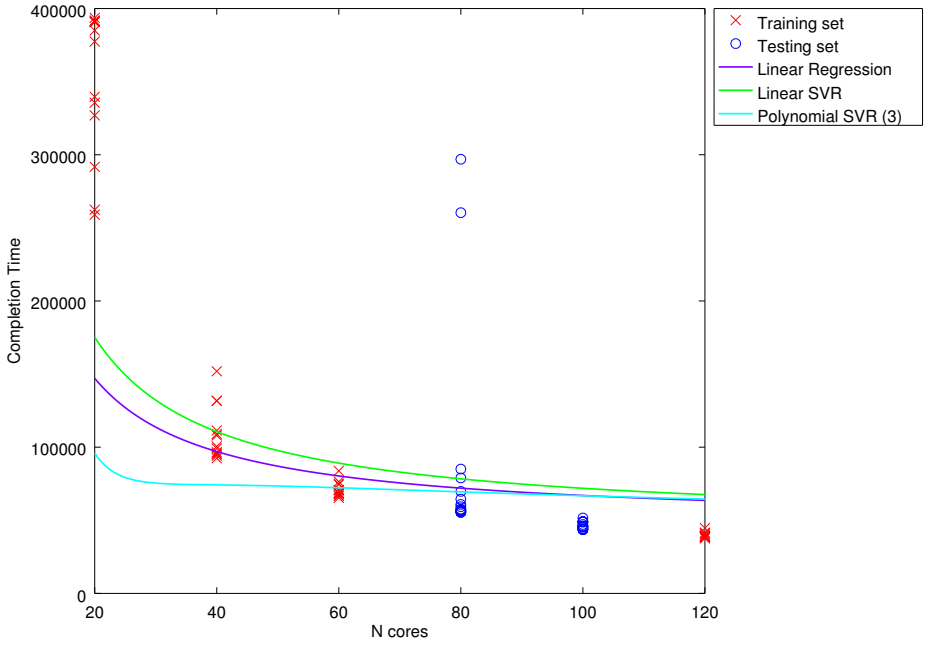

Figure 3: Completion time vs ncores for R1 (250GB), testing on 80, 100 cores

1.1.4 Query R1, Datasize 500GB — testing on 40, 60 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1903	-2.2417	39498	0.1006
Linear SVR	0.2192	0.1299	40157	0.1030
Polynomial SVR (2)	1.2736	0.0416	291006	0.7501
Polynomial SVR (3)	0.9263	0.2090	212552	0.5464
Polynomial SVR (4)	1.1752	0.3214	269171	0.6911
Polynomial SVR (6)	1.1382	0.0491	260931	0.6705
Gaussian SVR	0.2666	0.1701	52823	0.1382

Table 4: Results for R1 (500GB), testing on 40, 60 cores

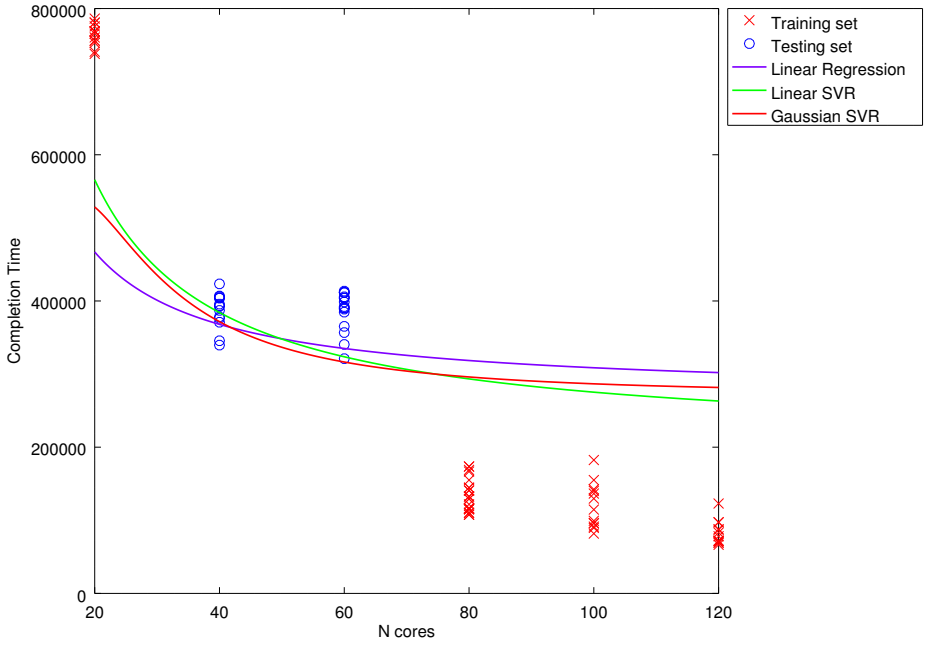

Figure 4: Completion time vs ncores for R1 (500GB), testing on 40, 60 cores

1.1.5 Query R1, Datasize 500GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0799	0.9778	12920	0.0462
Linear SVR	0.1174	0.9965	23356	0.1076
Polynomial SVR (2)	0.5225	0.1257	94883	0.5001
Polynomial SVR (3)	0.2428	0.8643	46726	0.2270
Polynomial SVR (4)	0.4159	0.4732	85924	0.4789
Polynomial SVR (6)	0.4428	0.3205	86760	0.4111
Gaussian SVR	0.2074	0.8632	40022	0.2116

Table 5: Results for R1 (500GB), testing on 60, 80 cores

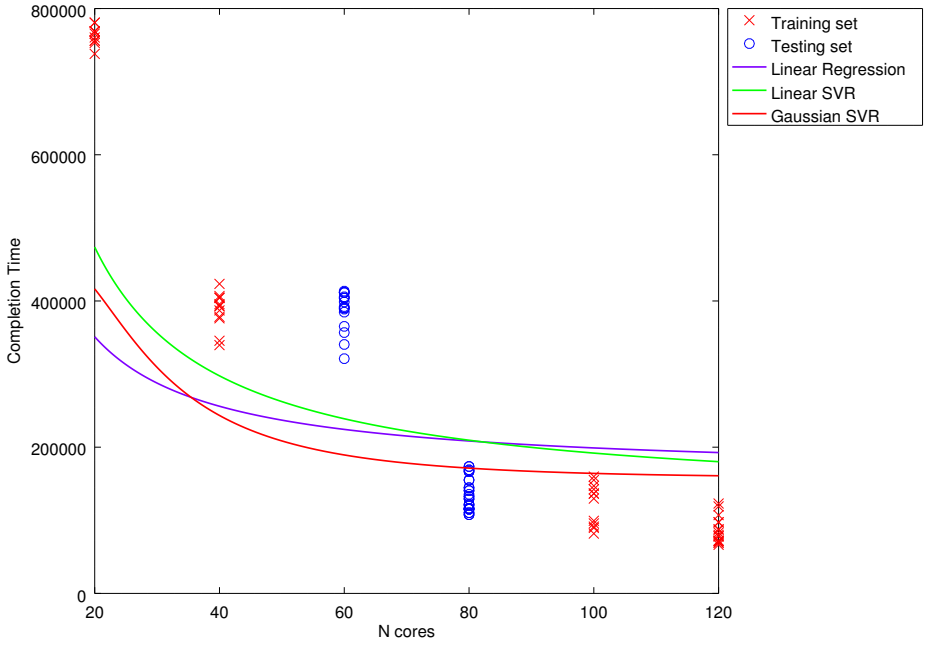

Figure 5: Completion time vs ncores for R1 (500GB), testing on 60, 80 cores

1.1.6 Query R1, Datasize 500GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0536	0.7523	10421	0.0812
Linear SVR	0.0586	0.8851	10865	0.0869
Polynomial SVR (2)	0.8840	0.0005	192135	1.5399
Polynomial SVR (3)	0.4204	0.1567	87786	0.6671
Polynomial SVR (4)	0.8494	0.0037	190904	1.5198
Polynomial SVR (6)	0.9483	0.0208	215729	1.7288
Gaussian SVR	0.2496	0.6860	46556	0.3363

Table 6: Results for R1 (500GB), testing on 80, 100 cores

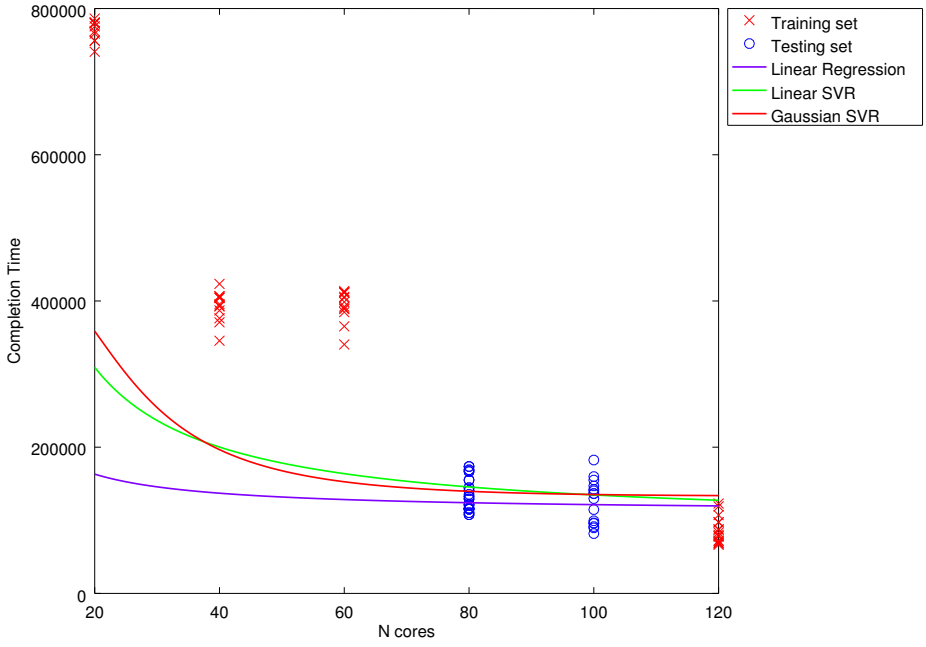

Figure 6: Completion time vs ncores for R1 (500GB), testing on 80, 100 cores

1.1.7 Query R1, Datasize 750GB — testing on 40, 60 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0327	0.9879	10529	0.0225
Linear SVR	0.0433	0.9837	11677	0.0246
Polynomial SVR (2)	0.9726	0.4374	311708	0.5999
Polynomial SVR (3)	0.5907	0.8986	179577	0.3341
Polynomial SVR (4)	0.7653	0.3657	239606	0.4538
Polynomial SVR (6)	0.7272	0.2950	226743	0.4283
Gaussian SVR	0.1548	0.8956	44369	0.0926

Table 7: Results for R1 (750GB), testing on 40, 60 cores

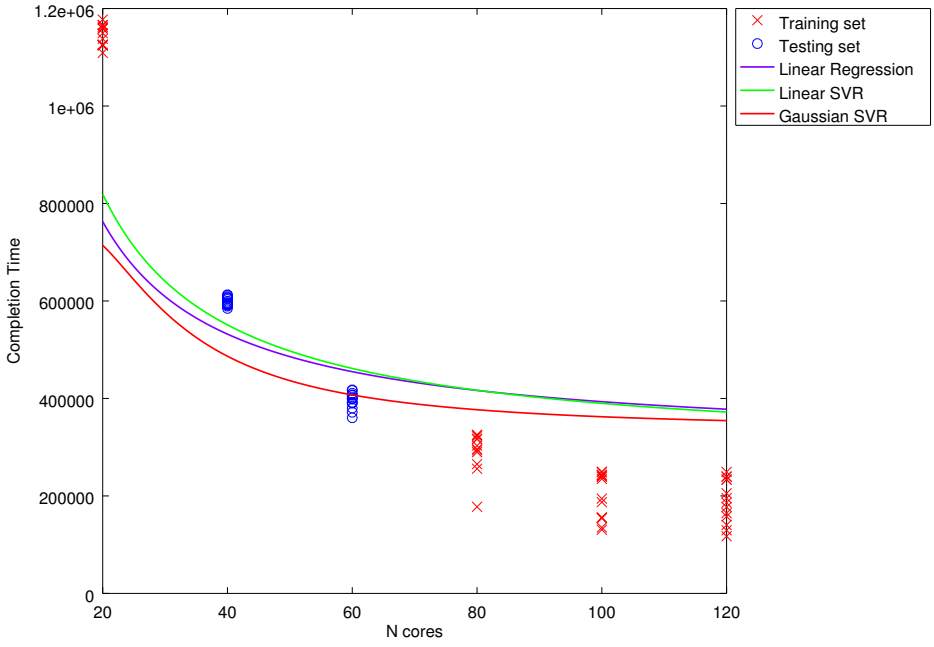

Figure 7: Completion time vs ncores for R1 (750GB), testing on 40, 60 cores

1.1.8 Query R1, Datasize 750GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0416	0.9594	10758	0.0348
Linear SVR	0.0576	0.9642	14924	0.0559
Polynomial SVR (2)	0.7932	0.3334	207890	0.6111
Polynomial SVR (3)	0.3994	0.6194	124462	0.3813
Polynomial SVR (4)	0.2774	0.4816	81163	0.2759
Polynomial SVR (6)	0.2541	0.4289	70791	0.2508
Gaussian SVR	0.1559	0.7164	43751	0.1332

Table 8: Results for R1 (750GB), testing on 60, 80 cores

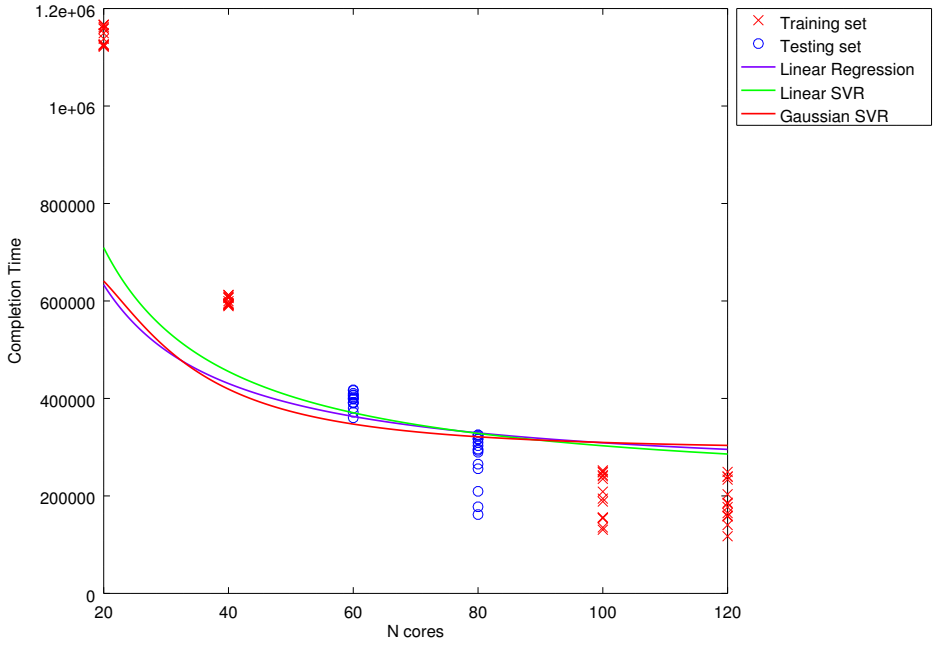

Figure 8: Completion time vs ncores for R1 (750GB), testing on 60, 80 cores

1.1.9 Query R1, Datasize 750GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0593	0.8861	19015	0.0898
Linear SVR	0.0801	0.8741	24070	0.1138
Polynomial SVR (2)	0.5867	0.7621	171631	0.6813
Polynomial SVR (3)	0.4020	0.2808	98863	0.4087
Polynomial SVR (4)	0.5525	0.2138	142699	0.6127
Polynomial SVR (6)	1.0243	0.0539	199660	0.8836
Gaussian SVR	0.1442	0.8211	40951	0.1713

Table 9: Results for R1 (750GB), testing on 80, 100 cores

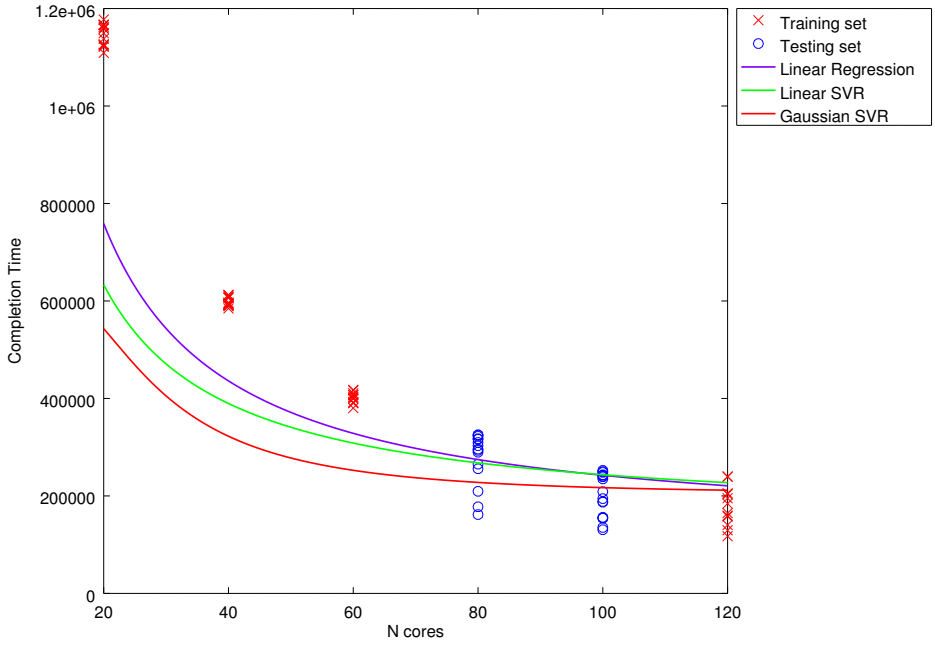

Figure 9: Completion time vs ncores for R1 (750GB), testing on 80, 100 cores

1.1.10 Query R1, Datasize 1000GB — testing on 40, 60 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0951	0.9471	23839	0.0262
Linear SVR	0.1399	0.9538	44642	0.0558
Polynomial SVR (2)	1.2992	0.2349	490795	0.6120
Polynomial SVR (3)	0.8388	0.3771	323627	0.4046
Polynomial SVR (4)	1.0080	0.1989	387625	0.4837
Polynomial SVR (6)	0.9506	0.1065	367190	0.4590
Gaussian SVR	0.4305	0.8307	140899	0.1597

Table 10: Results for R1 (1000GB), testing on 40, 60 cores

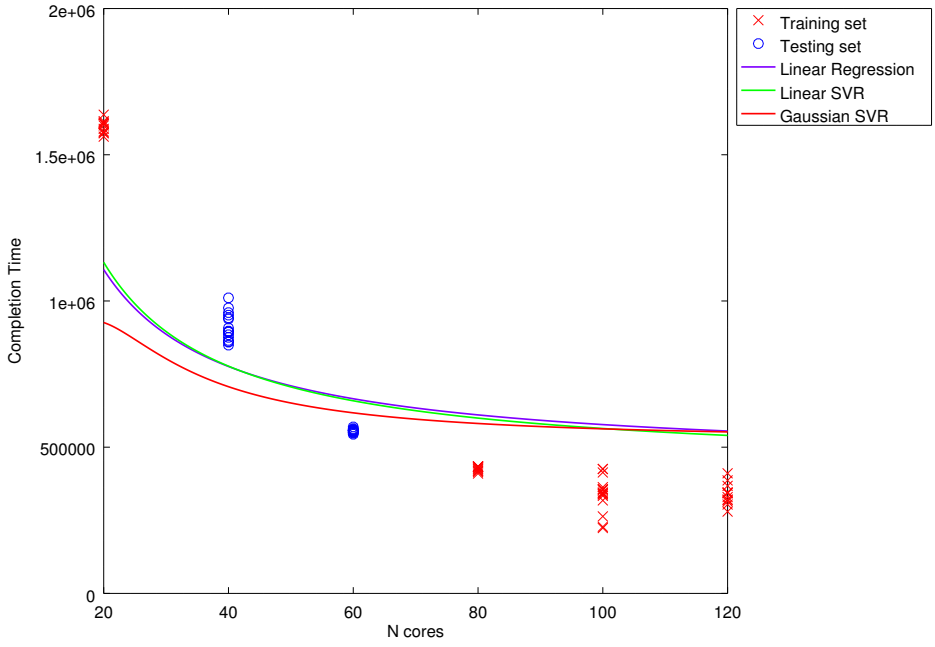

Figure 10: Completion time vs ncores for R1 (1000GB), testing on 40, 60 cores

1.1.11 Query R1, Datasize 1000GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0278	0.9673	11135	0.0236
Linear SVR	0.0247	0.9754	7390	0.0163
Polynomial SVR (2)	0.2691	0.2930	98133	0.1943
Polynomial SVR (3)	0.1516	0.5321	49358	0.1143
Polynomial SVR (4)	0.1618	0.0041	66228	0.1421
Polynomial SVR (6)	0.1550	0.0000	65046	0.1385
Gaussian SVR	0.0729	0.8406	24148	0.0522

Table 11: Results for R1 (1000GB), testing on 60, 80 cores

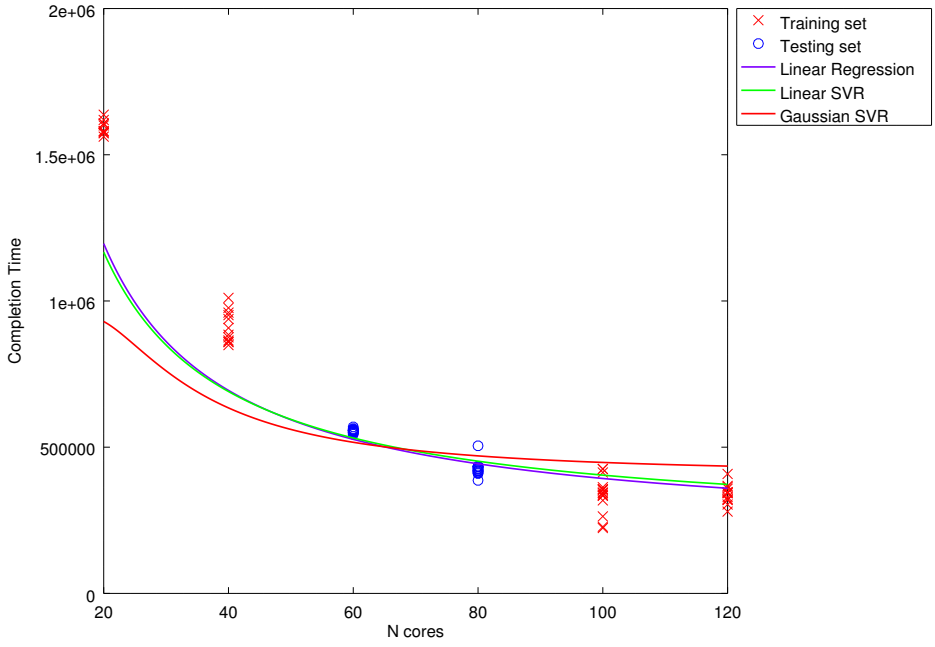

Figure 11: Completion time vs ncores for R1 (1000GB), testing on 60, 80 cores

1.1.12 Query R1, Datasize 1000GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0907	0.5955	28350	0.0771
Linear SVR	0.0999	0.9322	40918	0.1108
Polynomial SVR (2)	0.5146	0.2478	207576	0.5627
Polynomial SVR (3)	1.3142	0.5123	282595	1.0006
Polynomial SVR (4)	2.3210	0.3771	373908	1.4251
Polynomial SVR (6)	15.2811	0.3407	1623343	6.9967
Gaussian SVR	0.4060	0.1775	98263	0.3352

Table 12: Results for R1 (1000GB), testing on 80, 100 cores

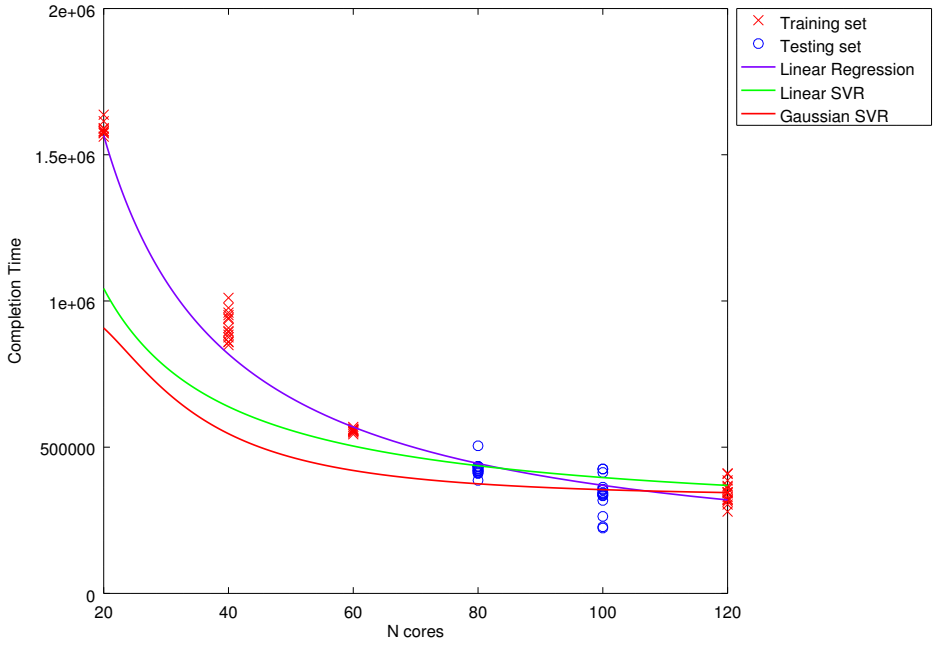

Figure 12: Completion time vs ncores for R1 (1000GB), testing on 80, 100 cores

1.2 Query R2

1.2.1 Query R2, Datasize 250GB — testing on 40, 60 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3588	0.8065	1190	0.0142
Linear SVR	0.4111	0.7866	1364	0.0162
Polynomial SVR (2)	1.0337	0.0190	3403	0.0396
Polynomial SVR (3)	0.6332	0.5734	2208	0.0260
Polynomial SVR (4)	0.8935	0.0070	2895	0.0339
Polynomial SVR (6)	0.8504	0.0233	2866	0.0337
Gaussian SVR	0.4638	0.7387	1518	0.0180

Table 13: Results for R2 (250GB), testing on 40, 60 cores

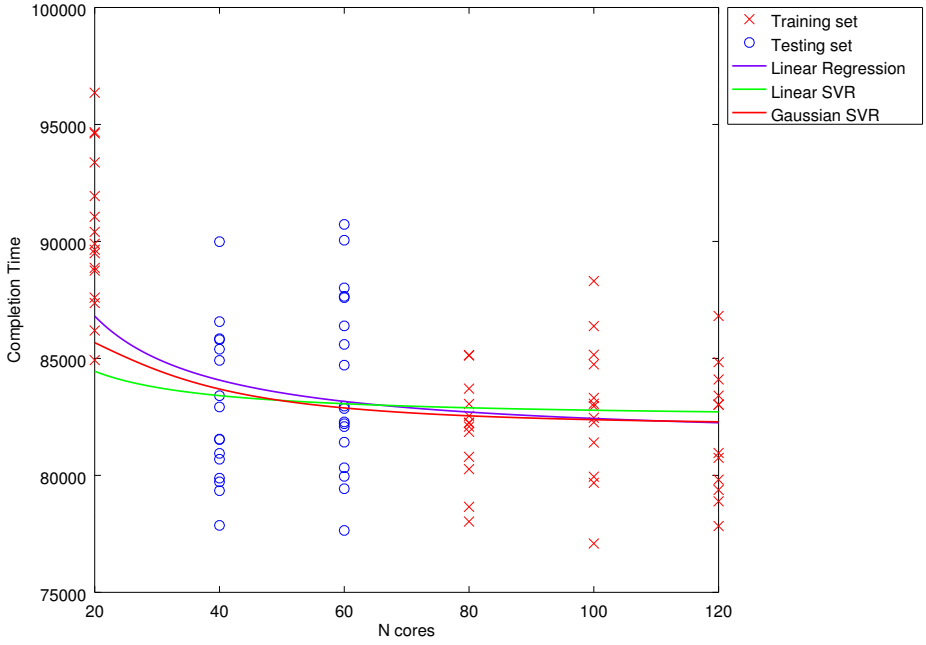

Figure 13: Completion time vs ncores for R2 (250GB), testing on 40, 60 cores

1.2.2 Query R2, Datasize 250GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3609	0.7655	1115	0.0132
Linear SVR	0.3796	0.7718	1230	0.0146
Polynomial SVR (2)	0.8043	0.0014	2574	0.0305
Polynomial SVR (3)	0.5799	0.5029	1913	0.0227
Polynomial SVR (4)	0.7476	0.0037	2463	0.0293
Polynomial SVR (6)	0.6737	0.3858	2222	0.0264
Gaussian SVR	0.3660	0.7771	1035	0.0122

Table 14: Results for R2 (250GB), testing on 60, 80 cores

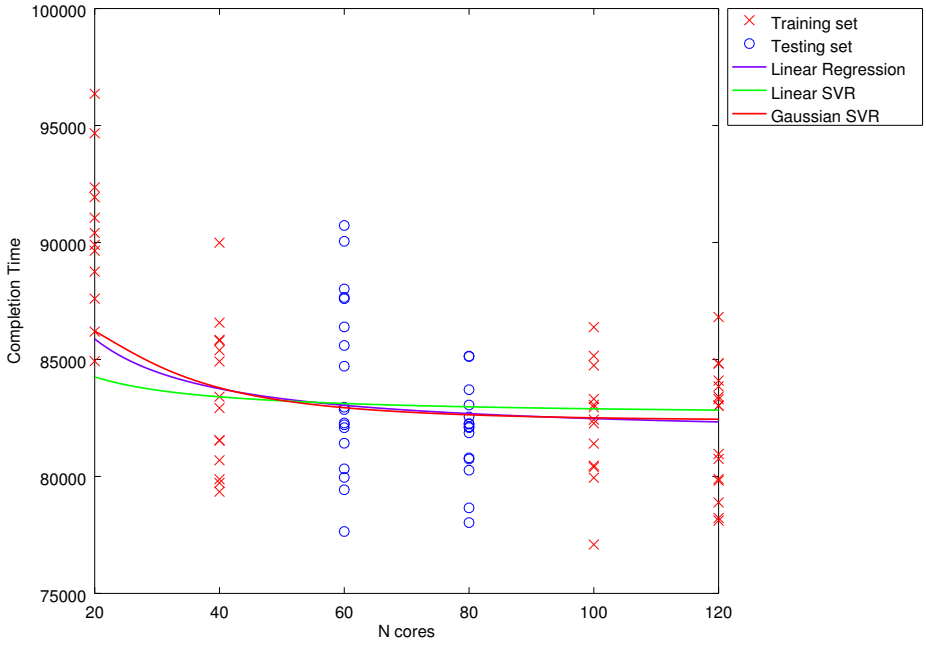

Figure 14: Completion time vs ncores for R2 (250GB), testing on 60, 80 cores

1.2.3 Query R2, Datasize 250GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2430	0.8276	851	0.0104
Linear SVR	0.2495	0.8280	864	0.0105
Polynomial SVR (2)	0.7199	0.0453	2516	0.0311
Polynomial SVR (3)	0.4766	0.5073	1659	0.0204
Polynomial SVR (4)	0.5516	0.3176	1841	0.0228
Polynomial SVR (6)	0.5899	0.3820	1976	0.0244
Gaussian SVR	0.3237	0.7679	1171	0.0143

Table 15: Results for R2 (250GB), testing on 80, 100 cores

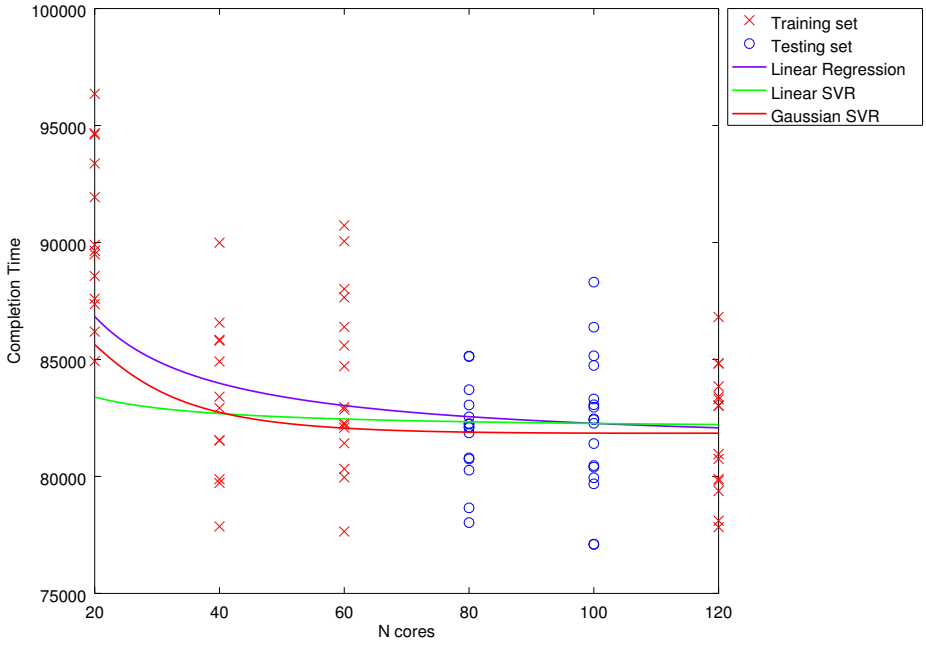

Figure 15: Completion time vs ncores for R2 (250GB), testing on 80, 100 cores

1.2.4 Query R2, Datasize 500GB — testing on 40, 60 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3423	0.9340	1140	0.0135
Linear SVR	0.7914	0.8817	3264	0.0381
Polynomial SVR (2)	2.9169	0.3151	10612	0.1218
Polynomial SVR (3)	1.1540	0.5111	5122	0.0612
Polynomial SVR (4)	3.1265	0.2306	10943	0.1260
Polynomial SVR (6)	1.3924	0.0000	7191	0.0891
Gaussian SVR	1.2324	0.7516	5327	0.0623

Table 16: Results for R2 (500GB), testing on 40, 60 cores

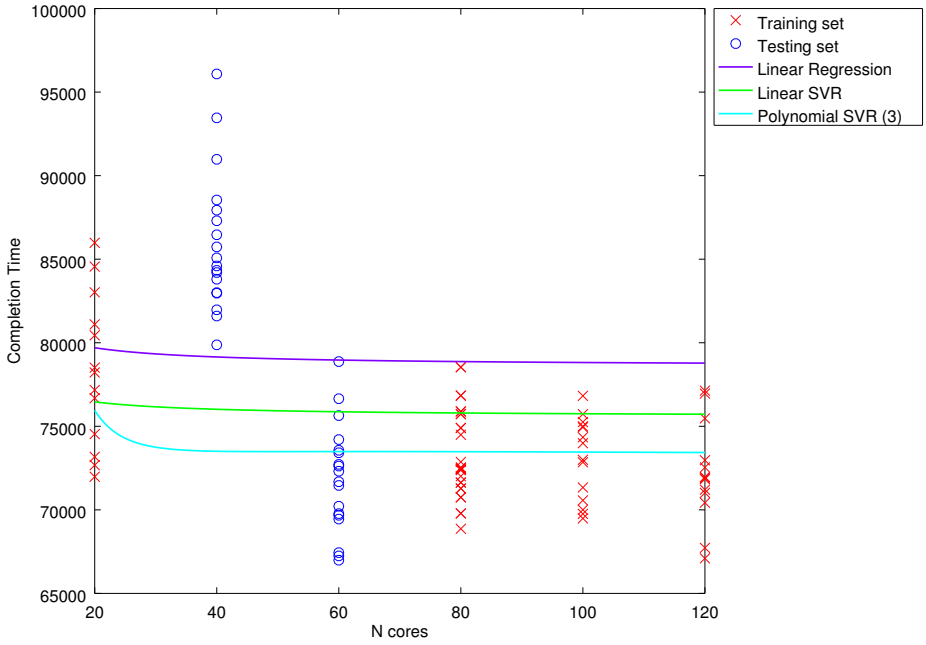

Figure 16: Completion time vs ncores for R2 (500GB), testing on 40, 60 cores

1.2.5 Query R2, Datasize 500GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1279	0.9335	685	0.0094
Linear SVR	0.0902	0.9825	438	0.0061
Polynomial SVR (2)	0.7196	0.7616	3454	0.0481
Polynomial SVR (3)	0.4301	0.7028	2092	0.0296
Polynomial SVR (4)	0.6658	0.2305	3450	0.0484
Polynomial SVR (6)	0.6906	0.1489	3459	0.0485
Gaussian SVR	0.2260	0.9010	985	0.0138

Table 17: Results for R2 (500GB), testing on 60, 80 cores

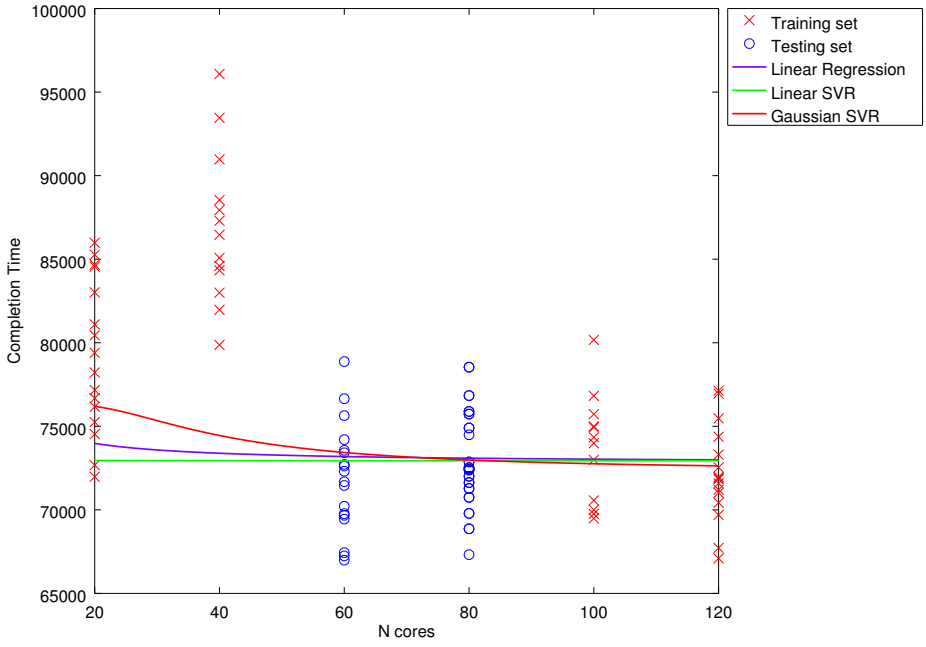

Figure 17: Completion time vs ncores for R2 (500GB), testing on 60, 80 cores

1.2.6 Query R2, Datasize 500GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0976	0.9588	454	0.0061
Linear SVR	0.0666	0.9846	302	0.0041
Polynomial SVR (2)	0.6795	0.7641	3273	0.0447
Polynomial SVR (3)	0.4372	0.7104	2229	0.0311
Polynomial SVR (4)	0.5998	0.0000	2509	0.0342
Polynomial SVR (6)	0.7580	0.0015	3918	0.0545
Gaussian SVR	0.1257	0.9580	566	0.0077

Table 18: Results for R2 (500GB), testing on 80, 100 cores

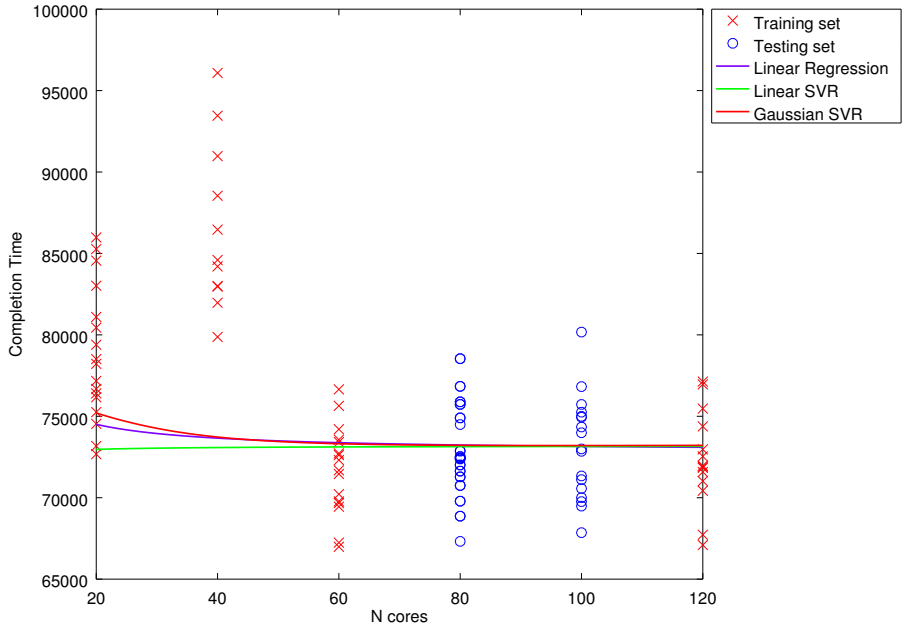

Figure 18: Completion time vs ncores for R2 (500GB), testing on 80, 100 cores

1.2.7 Query R2, Datasize 750GB — testing on 40, 60 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1798	0.9706	452	0.0057
Linear SVR	0.1759	0.9729	427	0.0054
Polynomial SVR (2)	1.4683	0.3643	3287	0.0394
Polynomial SVR (3)	0.6928	0.6389	1853	0.0227
Polynomial SVR (4)	2.1409	0.4196	3939	0.0469
Polynomial SVR (6)	1.8638	0.3868	3695	0.0441
Gaussian SVR	0.6016	0.7836	1320	0.0160

Table 19: Results for R2 (750GB), testing on 40, 60 cores

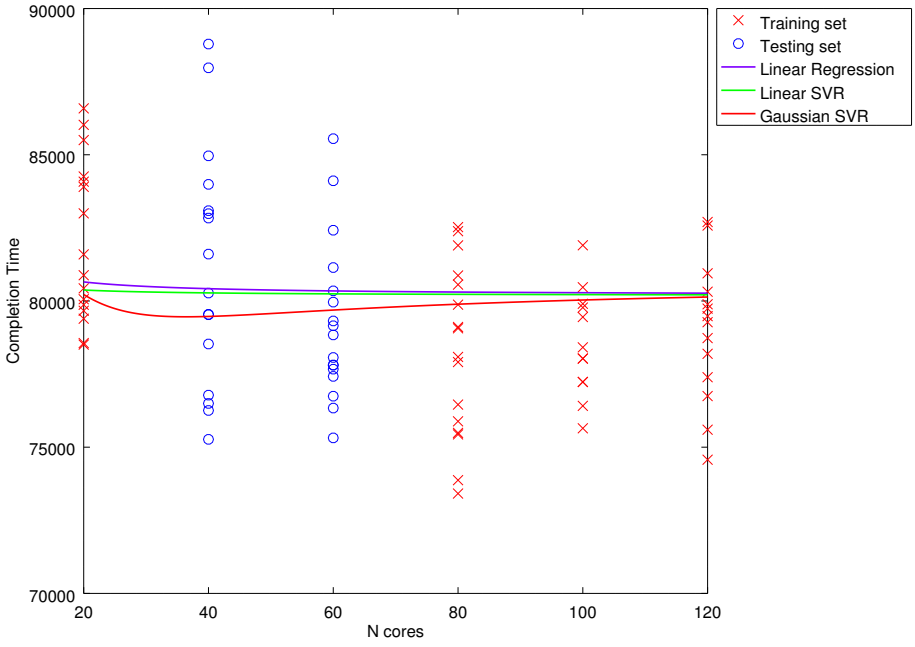

Figure 19: Completion time vs ncores for R2 (750GB), testing on 40, 60 cores

1.2.8 Query R2, Datasize 750GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1874	0.9488	521	0.0066
Linear SVR	0.2220	0.9433	562	0.0071
Polynomial SVR (2)	1.0127	0.1650	2622	0.0337
Polynomial SVR (3)	0.5024	0.7663	1397	0.0178
Polynomial SVR (4)	0.8482	0.0096	2273	0.0290
Polynomial SVR (6)	0.7515	0.3102	2077	0.0265
Gaussian SVR	0.1764	0.9553	489	0.0062

Table 20: Results for R2 (750GB), testing on 60, 80 cores

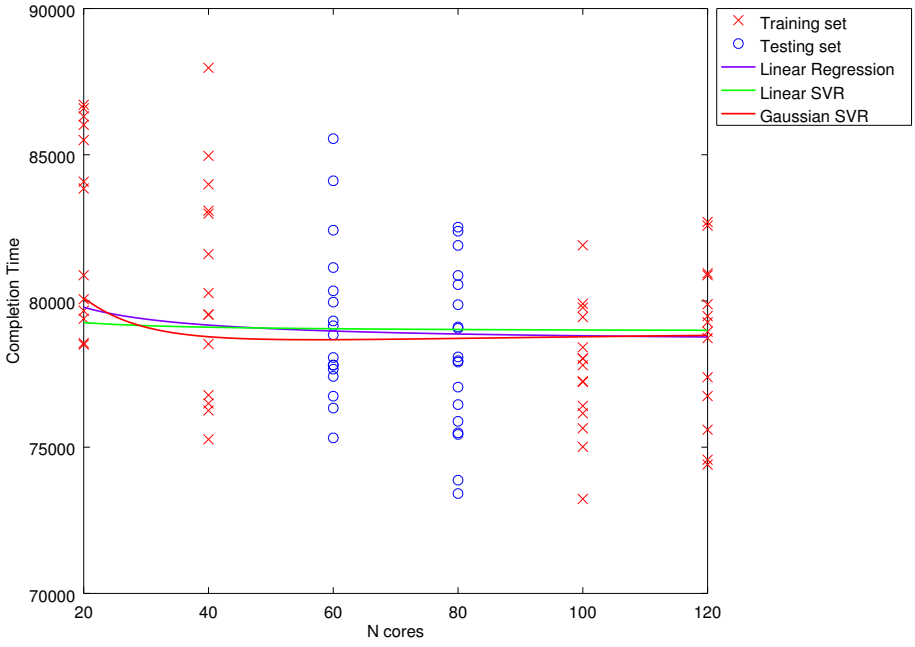

Figure 20: Completion time vs ncores for R2 (750GB), testing on 60, 80 cores

1.2.9 Query R2, Datasize 750GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1957	0.9297	485	0.0062
Linear SVR	0.2040	0.9300	505	0.0064
Polynomial SVR (2)	0.9494	0.3778	2476	0.0323
Polynomial SVR (3)	0.4973	0.6884	1426	0.0184
Polynomial SVR (4)	0.8522	0.0667	2285	0.0298
Polynomial SVR (6)	0.8079	0.5461	2233	0.0291
Gaussian SVR	0.3731	0.7650	803	0.0105

Table 21: Results for R2 (750GB), testing on 80, 100 cores

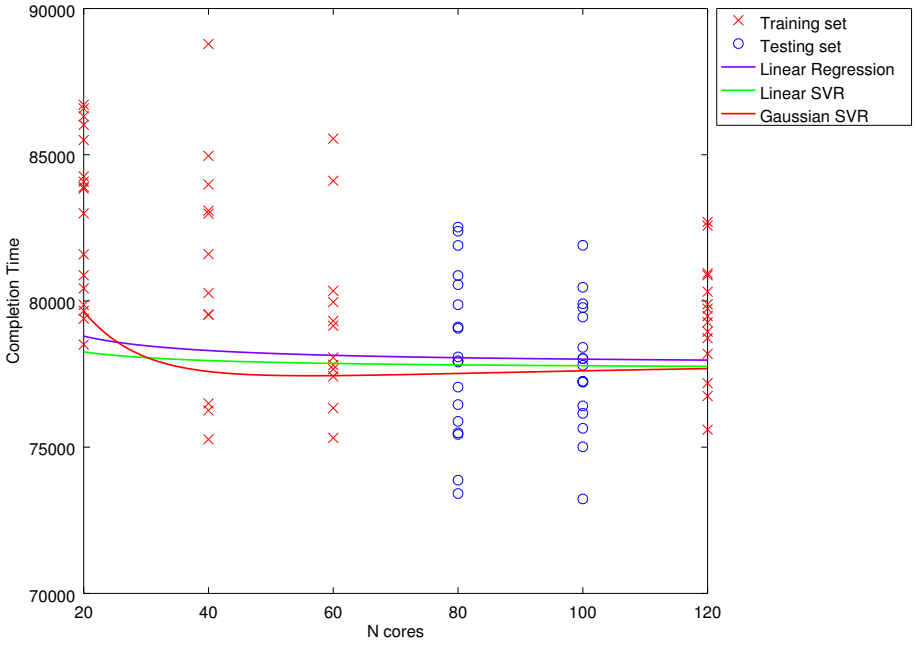

Figure 21: Completion time vs ncores for R2 (750GB), testing on 80, 100 cores

1.2.10 Query R2, Datasize 1000GB — testing on 40, 60 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2238	0.9098	259990	0.0736
Linear SVR	0.1465	0.9694	167434	0.0742
Polynomial SVR (2)	1.9452	0.7448	2702157	0.8153
Polynomial SVR (3)	0.6281	0.7242	899148	0.3030
Polynomial SVR (4)	1.9521	0.5170	2539061	0.7341
Polynomial SVR (6)	1.4676	0.3552	1951207	0.5654
Gaussian SVR	0.4864	0.8661	609021	0.1734

Table 22: Results for R2 (1000GB), testing on 40, 60 cores

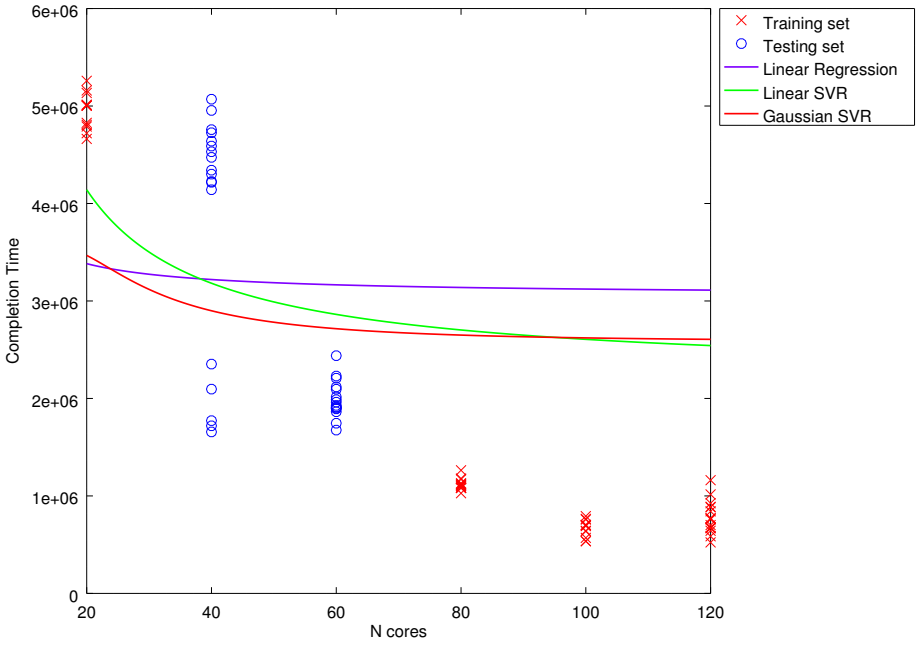

Figure 22: Completion time vs ncores for R2 (1000GB), testing on 40, 60 cores

1.2.11 Query R2, Datasize 1000GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1047	0.8471	167777	0.1105
Linear SVR	0.0191	0.9968	26918	0.0206
Polynomial SVR (2)	0.2549	0.3569	310668	0.1764
Polynomial SVR (3)	0.2433	0.4917	367294	0.2369
Polynomial SVR (4)	0.4475	0.8051	580356	0.3167
Polynomial SVR (6)	0.2902	0.0113	424739	0.2579
Gaussian SVR	0.1625	0.9547	193968	0.1084

Table 23: Results for R2 (1000GB), testing on 60, 80 cores

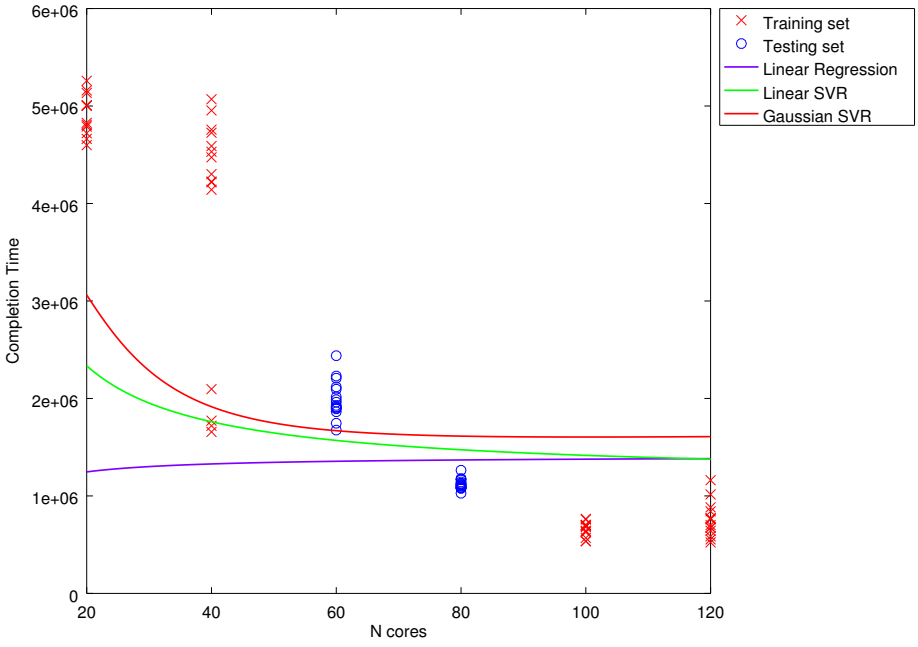

Figure 23: Completion time vs ncores for R2 (1000GB), testing on 60, 80 cores

1.2.12 Query R2, Datasize 1000GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0185	0.9821	27260	0.0337
Linear SVR	0.0245	0.9857	33055	0.0354
Polynomial SVR (2)	0.2137	0.9443	325346	0.3426
Polynomial SVR (3)	0.2305	0.9815	371215	0.4037
Polynomial SVR (4)	0.3273	0.9516	539273	0.6010
Polynomial SVR (6)	0.4632	0.7098	760532	0.9738
Gaussian SVR	0.0430	0.9074	56926	0.0726

Table 24: Results for R2 (1000GB), testing on 80, 100 cores

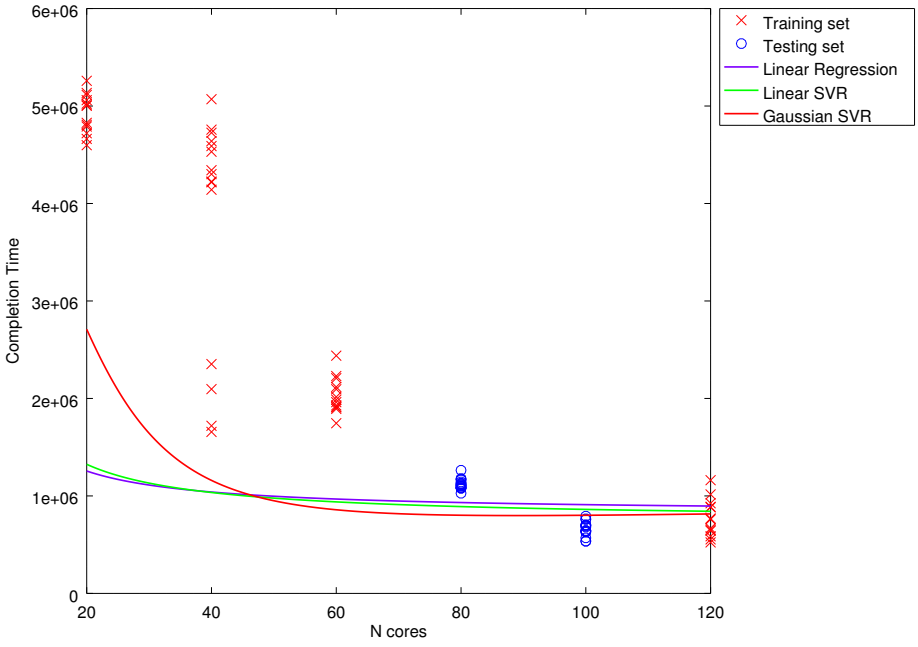

Figure 24: Completion time vs ncores for R2 (1000GB), testing on 80, 100 cores

1.3 Query R3

1.3.1 Query R3, Datasize 250GB — testing on 40, 60 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0762	0.9646	19812	0.0442
Linear SVR	0.1664	0.9926	47142	0.1116
Polynomial SVR (2)	1.0154	0.0233	290567	0.6758
Polynomial SVR (3)	0.7307	0.2822	193900	0.4176
Polynomial SVR (4)	0.8772	0.6843	239466	0.5301
Polynomial SVR (6)	0.8458	0.6479	230431	0.5086
Gaussian SVR	0.3187	0.9548	84376	0.1928

Table 25: Results for R3 (250GB), testing on 40, 60 cores

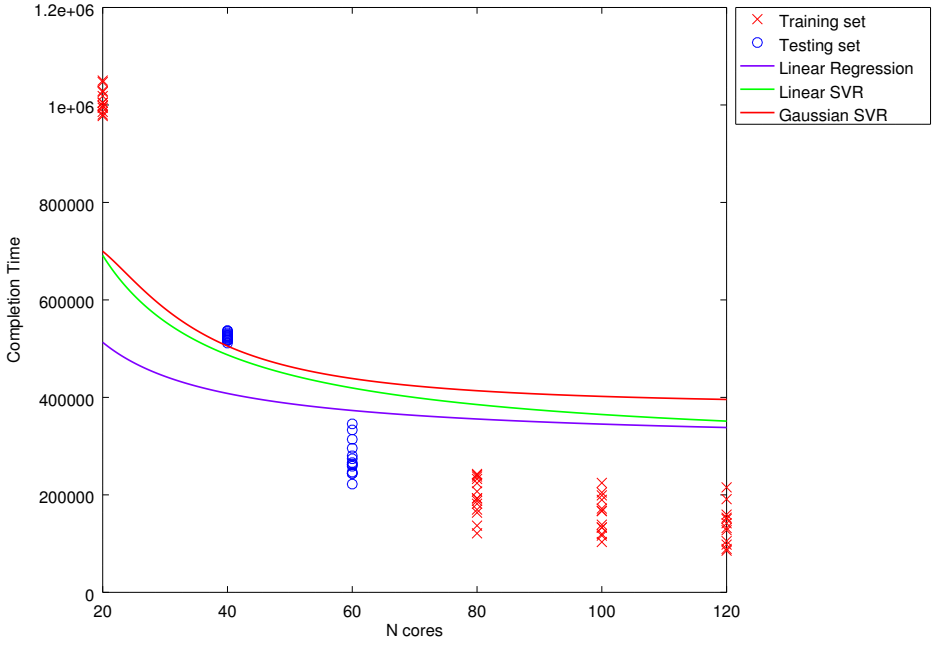

Figure 25: Completion time vs ncores for R3 (250GB), testing on 40, 60 cores

1.3.2 Query R3, Datasize 250GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0291	0.9728	7120	0.0283
Linear SVR	0.1062	0.9601	30148	0.1313
Polynomial SVR (2)	0.6215	0.4521	164153	0.6679
Polynomial SVR (3)	0.1136	0.6781	28855	0.1419
Polynomial SVR (4)	0.1914	0.0068	51645	0.2340
Polynomial SVR (6)	0.2003	0.1971	51538	0.2634
Gaussian SVR	0.1217	0.7286	28094	0.1236

Table 26: Results for R3 (250GB), testing on 60, 80 cores

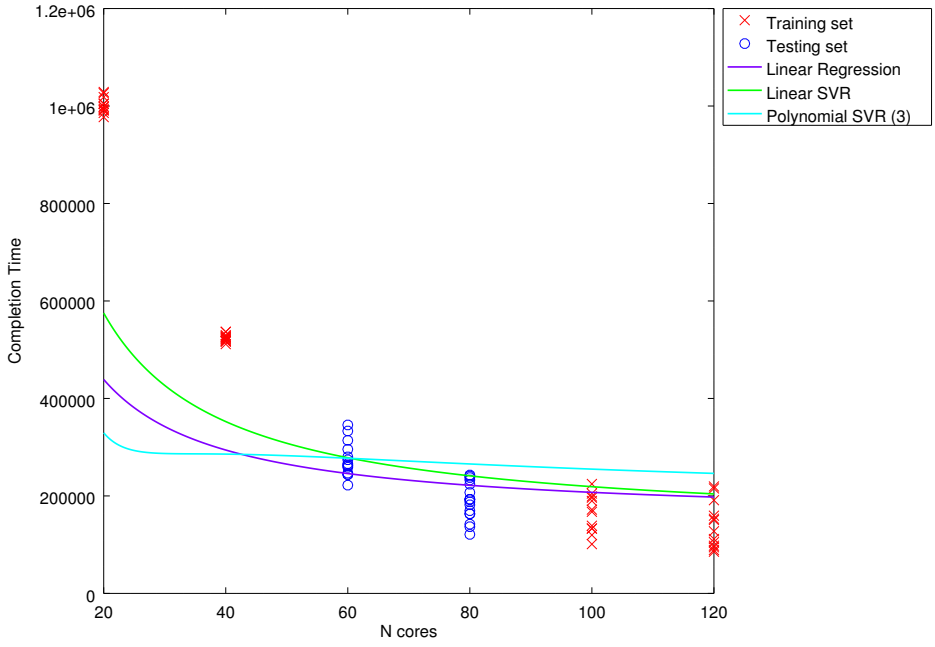

Figure 26: Completion time vs ncores for R3 (250GB), testing on 60, 80 cores

1.3.3 Query R3, Datasize 250GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0176	0.9824	3990	0.0213
Linear SVR	0.0467	0.8918	12878	0.0797
Polynomial SVR (2)	0.3637	0.3632	103963	0.6392
Polynomial SVR (3)	0.1651	0.4134	43444	0.2636
Polynomial SVR (4)	0.2726	0.3638	75571	0.4626
Polynomial SVR (6)	0.3226	0.3446	90660	0.5557
Gaussian SVR	0.0961	0.7726	24484	0.1478

Table 27: Results for R3 (250GB), testing on 80, 100 cores

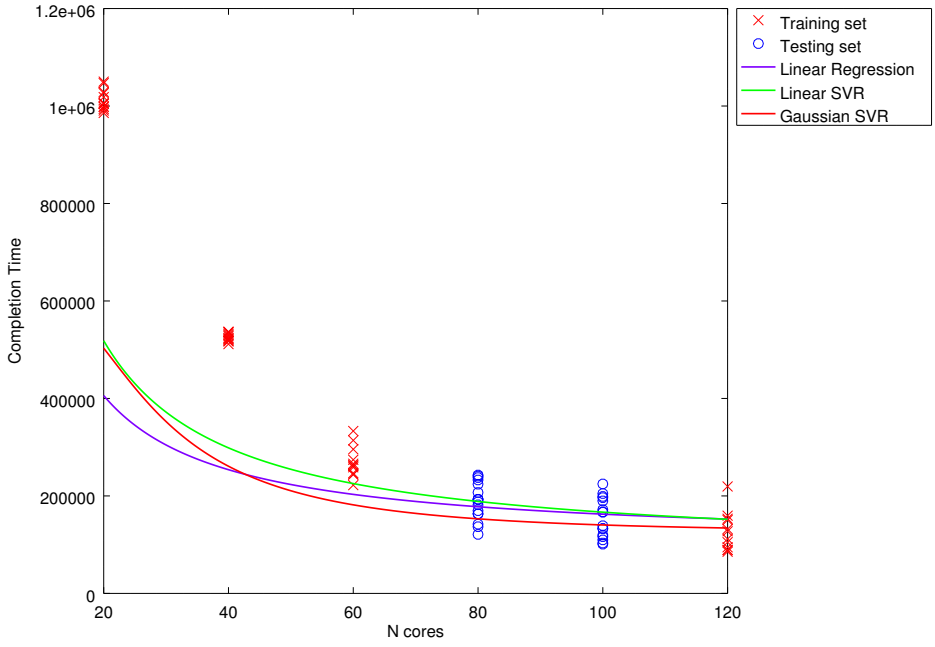

Figure 27: Completion time vs ncores for R3 (250GB), testing on 80, 100 cores

1.3.4 Query R3, Datasize 500GB — testing on 40, 60 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3279	-215.7546	135158	0.1313
Linear SVR	0.1905	0.0460	83075	0.0806
Polynomial SVR (2)	1.1115	0.0099	616532	0.5976
Polynomial SVR (3)	0.8999	0.1187	500084	0.4848
Polynomial SVR (4)	0.9946	0.0050	552796	0.5359
Polynomial SVR (6)	0.9908	0.0050	550734	0.5339
Gaussian SVR	0.4366	0.0171	225259	0.2181

Table 28: Results for R3 (500GB), testing on 40, 60 cores

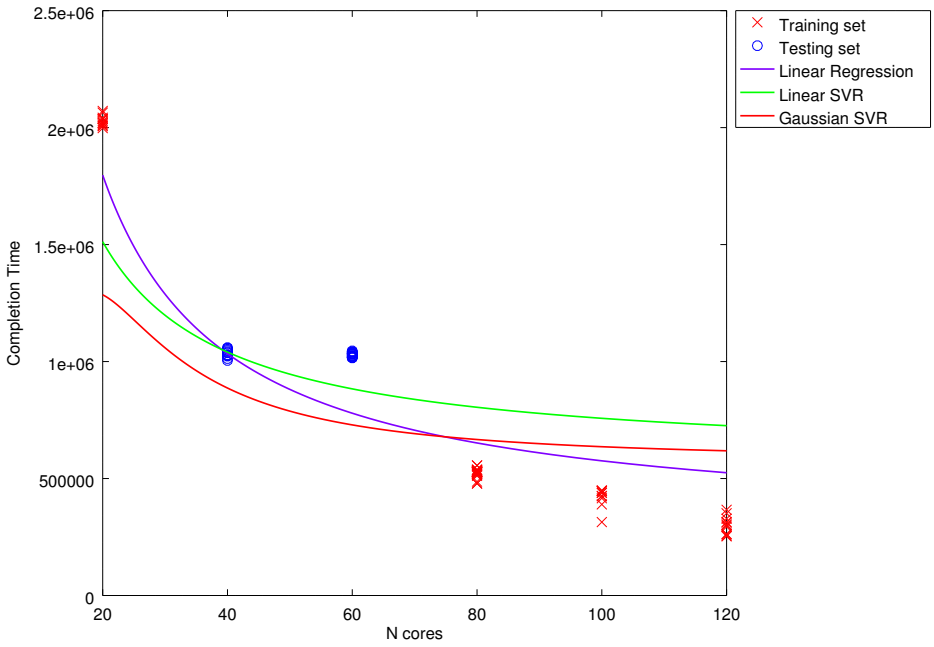

Figure 28: Completion time vs ncores for R3 (500GB), testing on 40, 60 cores

1.3.5 Query R3, Datasize 500GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2085	0.7742	75349	0.0782
Linear SVR	0.1593	0.9893	71627	0.0898
Polynomial SVR (2)	0.5283	0.0284	276187	0.4555
Polynomial SVR (3)	0.4090	0.2941	203839	0.3513
Polynomial SVR (4)	0.8084	0.0000	330135	0.6131
Polynomial SVR (6)	0.6143	0.0184	301580	0.5269
Gaussian SVR	0.2642	0.6436	116089	0.1744

Table 29: Results for R3 (500GB), testing on 60, 80 cores

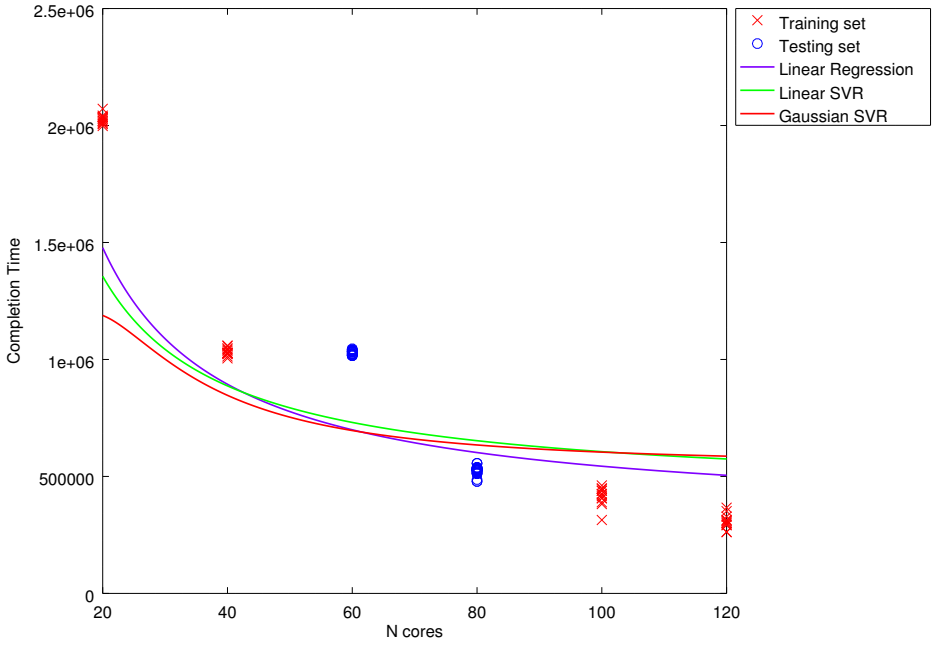

Figure 29: Completion time vs ncores for R3 (500GB), testing on 60, 80 cores

1.3.6 Query R3, Datasize 500GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0514	0.7602	25149	0.0504
Linear SVR	0.1586	0.9268	85071	0.1739
Polynomial SVR (2)	1.1213	0.5499	599035	1.2236
Polynomial SVR (3)	0.7612	0.5805	411543	0.8450
Polynomial SVR (4)	0.9068	0.4613	493082	1.0183
Polynomial SVR (6)	0.9627	0.1780	505340	1.0542
Gaussian SVR	0.5018	0.6248	249892	0.5018

Table 30: Results for R3 (500GB), testing on 80, 100 cores

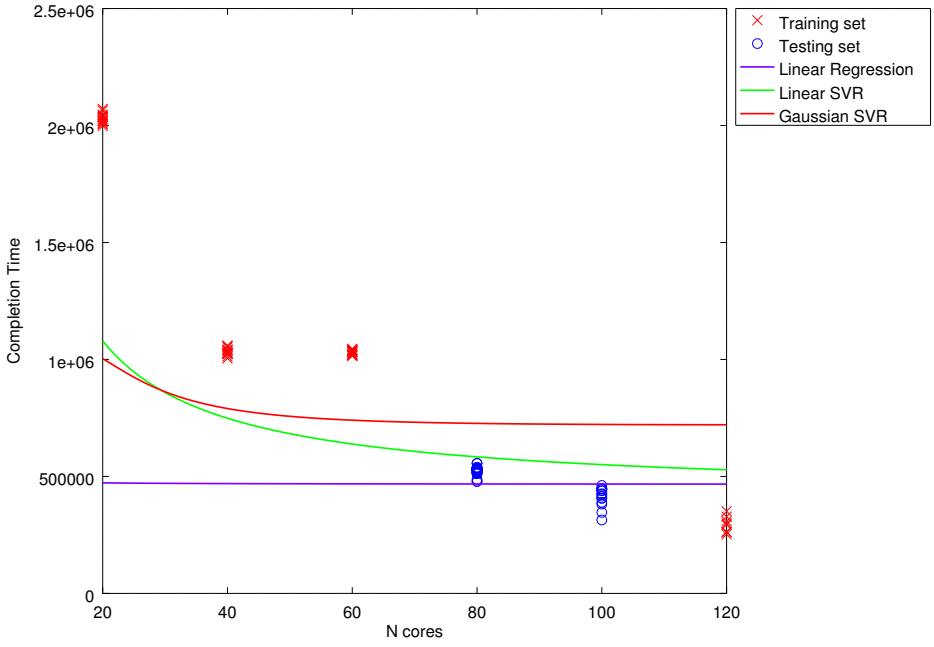

Figure 30: Completion time vs ncores for R3 (500GB), testing on 80, 100 cores

1.3.7 Query R3, Datasize 750GB — testing on 40, 60 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0258	0.9924	18906	0.0151
Linear SVR	0.2204	0.9372	174219	0.1360
Polynomial SVR (2)	0.8491	0.1628	637605	0.4720
Polynomial SVR (3)	0.6535	0.0045	494486	0.3646
Polynomial SVR (4)	0.7471	0.1440	585359	0.4400
Polynomial SVR (6)	0.7273	0.0051	563635	0.4208
Gaussian SVR	0.3253	0.4934	250650	0.1886

Table 31: Results for R3 (750GB), testing on 40, 60 cores

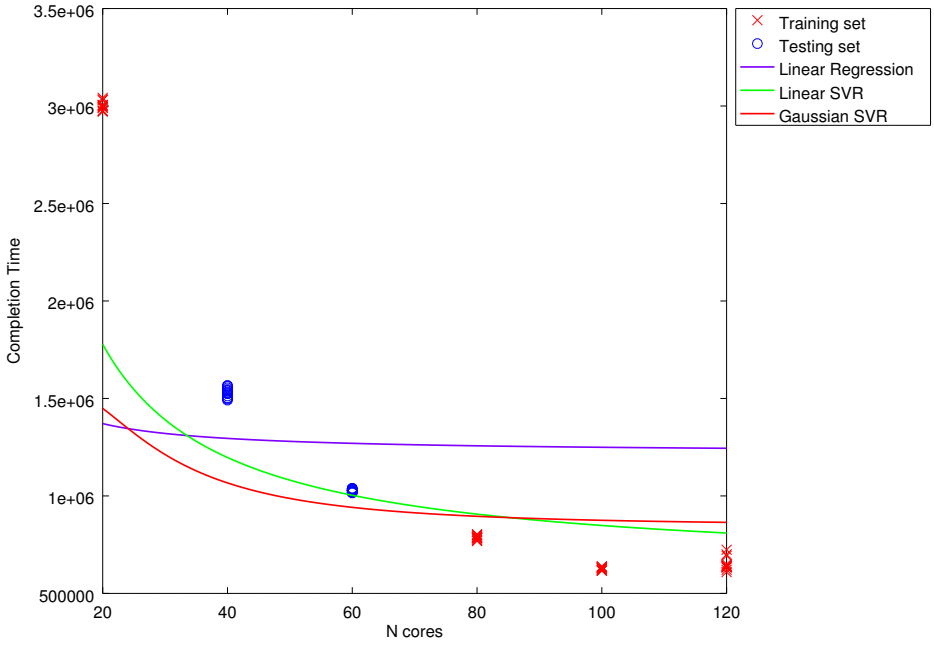

Figure 31: Completion time vs ncores for R3 (750GB), testing on 40, 60 cores

1.3.8 Query R3, Datasize 750GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0168	0.9861	13675	0.0149
Linear SVR	0.0403	0.9648	31659	0.0358
Polynomial SVR (2)	0.4169	0.2180	294885	0.3224
Polynomial SVR (3)	0.2978	0.3584	180981	0.2007
Polynomial SVR (4)	0.4573	0.1593	234477	0.2326
Polynomial SVR (6)	0.3423	0.1686	205913	0.2334
Gaussian SVR	0.1940	0.7040	104537	0.1084

Table 32: Results for R3 (750GB), testing on 60, 80 cores

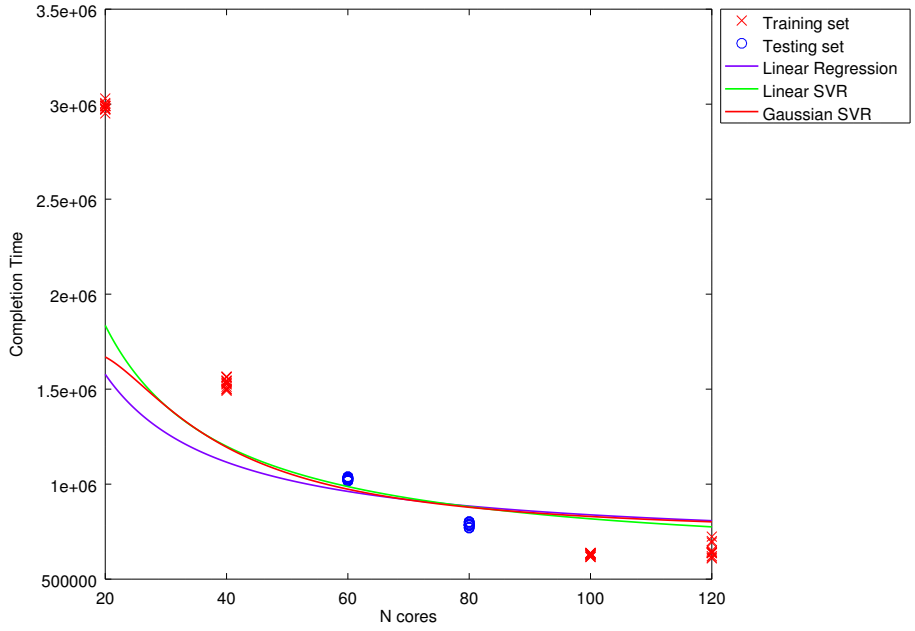

Figure 32: Completion time vs ncores for R3 (750GB), testing on 60, 80 cores

1.3.9 Query R3, Datasize 750GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0273	0.9153	22727	0.0326
Linear SVR	0.1061	0.9175	85459	0.1257
Polynomial SVR (2)	0.4300	0.4348	337071	0.4746
Polynomial SVR (3)	0.3126	0.4066	253679	0.3641
Polynomial SVR (4)	0.5151	0.2318	426915	0.6144
Polynomial SVR (6)	0.4808	0.1579	401164	0.5848
Gaussian SVR	0.1015	0.2202	64616	0.0996

Table 33: Results for R3 (750GB), testing on 80, 100 cores

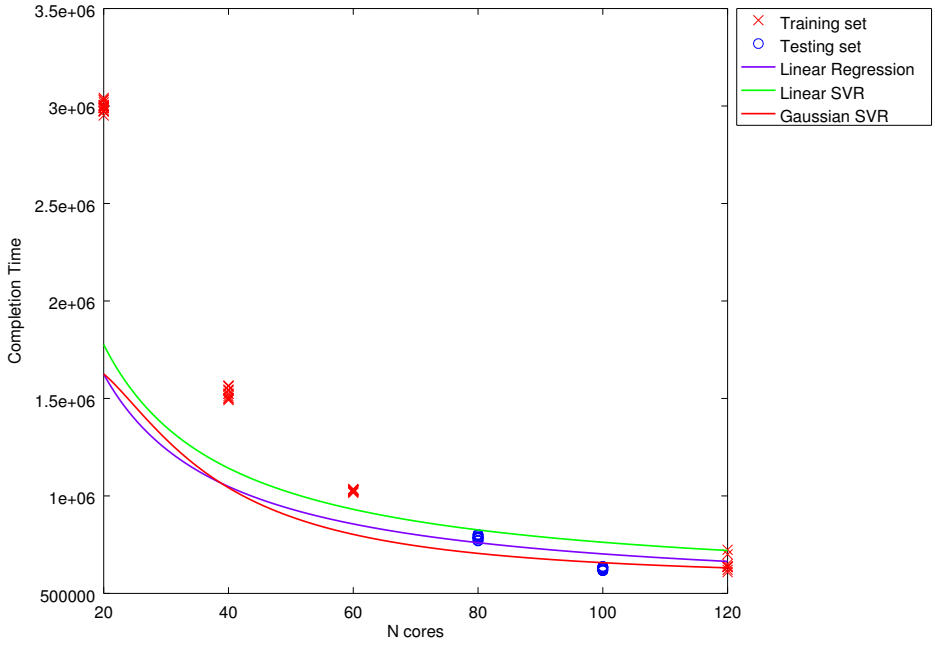

Figure 33: Completion time vs ncores for R3 (750GB), testing on 80, 100 cores

1.3.10 Query R3, Datasize 1000GB — testing on 40, 60 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0955	0.9193	85756	0.0455
Linear SVR	0.1046	0.9321	92534	0.0495
Polynomial SVR (2)	1.0220	0.3572	1107822	0.6207
Polynomial SVR (3)	0.5783	0.4947	608093	0.3355
Polynomial SVR (4)	0.7805	0.1606	821330	0.4511
Polynomial SVR (6)	0.7547	0.2127	793157	0.4349
Gaussian SVR	0.2407	0.8350	245122	0.1392

Table 34: Results for R3 (1000GB), testing on 40, 60 cores

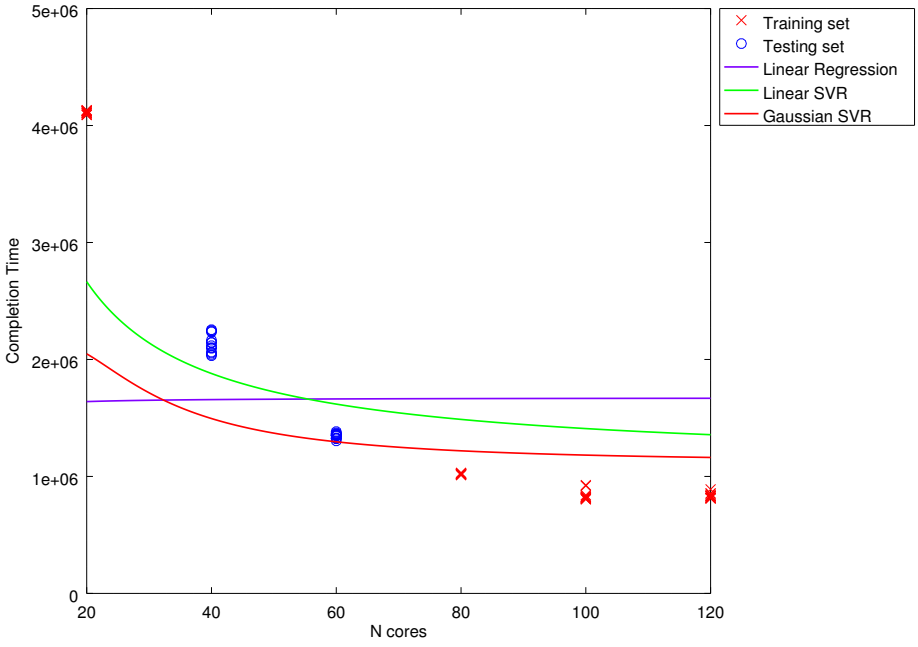

Figure 34: Completion time vs ncores for R3 (1000GB), testing on 40, 60 cores

1.3.11 Query R3, Datasize 1000GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0340	0.9396	35677	0.0289
Linear SVR	0.0517	0.9751	55717	0.0475
Polynomial SVR (2)	0.5918	0.4671	636845	0.5094
Polynomial SVR (3)	0.1245	0.7738	113290	0.1068
Polynomial SVR (4)	0.1491	0.6445	175289	0.1496
Polynomial SVR (6)	0.2710	0.6118	274633	0.2138
Gaussian SVR	0.2104	0.5312	223288	0.1833

Table 35: Results for R3 (1000GB), testing on 60, 80 cores

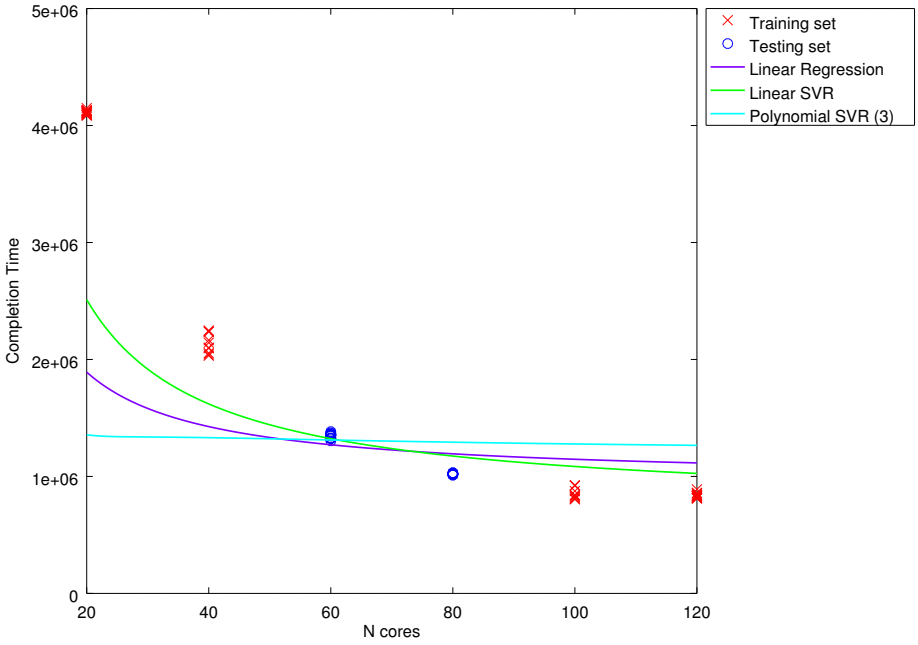

Figure 35: Completion time vs ncores for R3 (1000GB), testing on 60, 80 cores

1.3.12 Query R3, Datasize 1000GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.0445	0.6986	27965	0.0313
Linear SVR	0.0957	0.8070	104958	0.1168
Polynomial SVR (2)	0.3359	0.0320	375170	0.4190
Polynomial SVR (3)	0.1500	0.2305	153658	0.1710
Polynomial SVR (4)	0.2807	0.0033	316840	0.3581
Polynomial SVR (6)	0.2865	0.0284	324880	0.3667
Gaussian SVR	0.1786	0.0356	115302	0.1298

Table 36: Results for R3 (1000GB), testing on 80, 100 cores

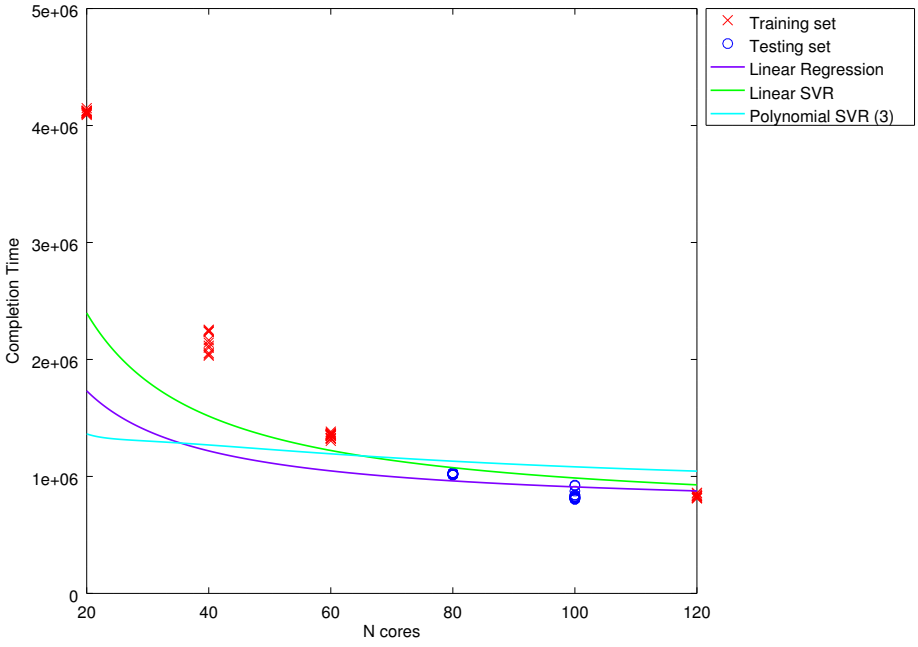

Figure 36: Completion time vs ncores for R3 (1000GB), testing on 80, 100 cores

1.4 Query R4

1.4.1 Query R4, Datasize 250GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3366	0.8540	12531	0.0528
Linear SVR	1.0866	0.7252	56099	0.2575
Polynomial SVR (2)	1.8978	0.6911	92636	0.4065
Polynomial SVR (3)	0.7723	0.7662	42157	0.2114
Polynomial SVR (4)	2.4925	0.6115	112859	0.4813
Polynomial SVR (6)	2.0438	0.5443	101427	0.4518
Gaussian SVR	1.3036	0.2434	64222	0.2830

Table 37: Results for R4 (250GB), testing on 60, 80 cores

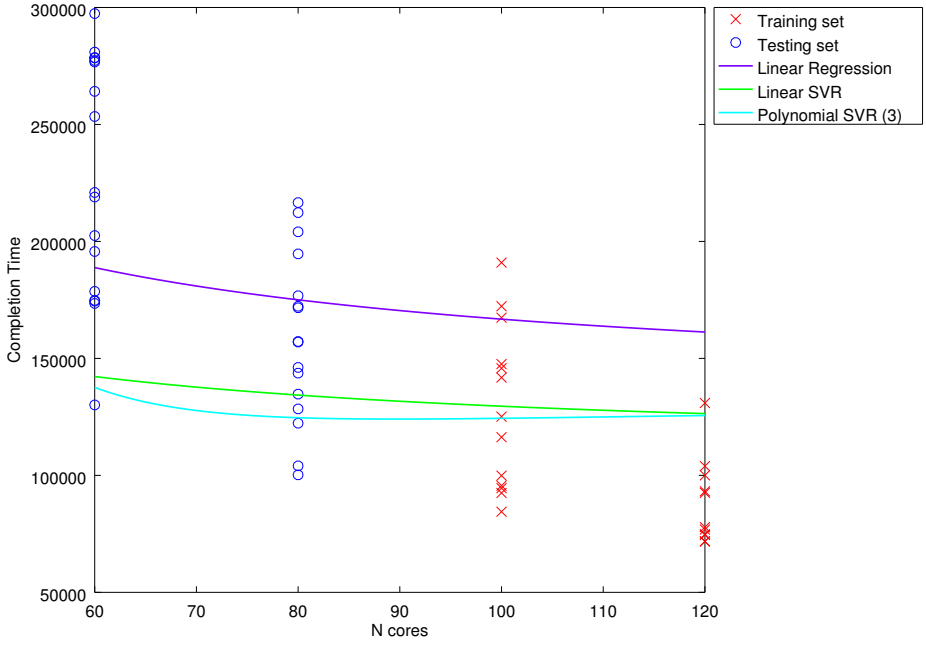

Figure 37: Completion time vs ncores for R4 (250GB), testing on 60, 80 cores

1.4.2 Query R4, Datasize 250GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1069	0.9662	4287	0.0268
Linear SVR	0.1142	0.9738	5113	0.0323
Polynomial SVR (2)	0.9712	0.0625	45169	0.2853
Polynomial SVR (3)	0.4437	0.7114	23362	0.1678
Polynomial SVR (4)	0.5712	0.0423	29166	0.2167
Polynomial SVR (6)	0.5653	0.1500	29339	0.2252
Gaussian SVR	0.2138	0.8793	10641	0.0693

Table 38: Results for R4 (250GB), testing on 80, 100 cores

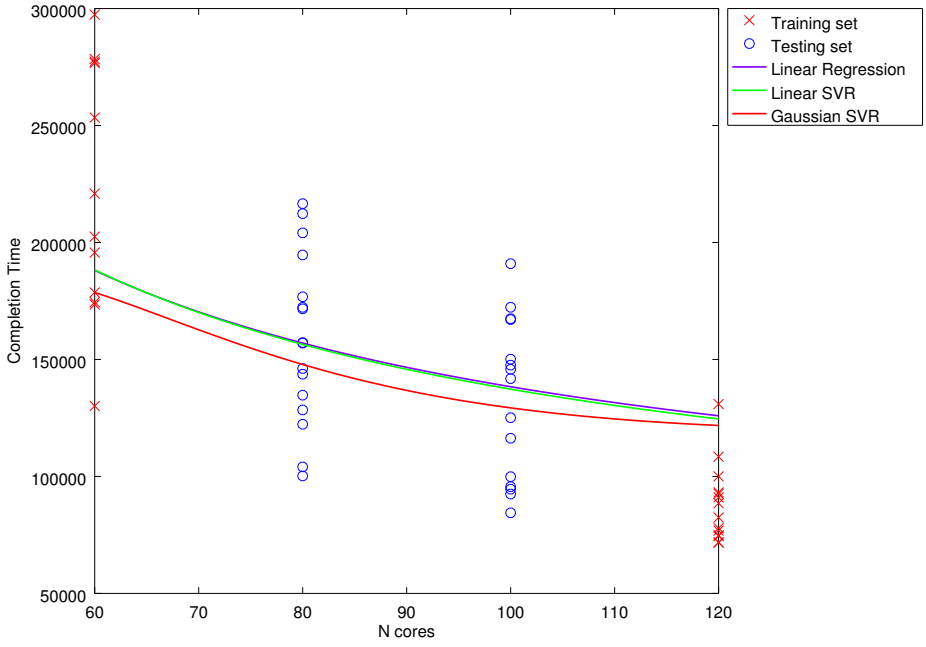

Figure 38: Completion time vs ncores for R4 (250GB), testing on 80, 100 cores

1.4.3 Query R4, Datasize 250GB — testing on 100, 120 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2352	0.8089	12457	0.1401
Linear SVR	0.2640	0.9909	13866	0.1580
Polynomial SVR (2)	1.5247	0.5057	80120	0.9107
Polynomial SVR (3)	0.4661	0.8374	25121	0.2348
Polynomial SVR (4)	1.7062	0.3001	86282	0.9834
Polynomial SVR (6)	1.0009	0.5137	55791	0.6203
Gaussian SVR	0.6979	0.6070	38144	0.4222

Table 39: Results for R4 (250GB), testing on 100, 120 cores

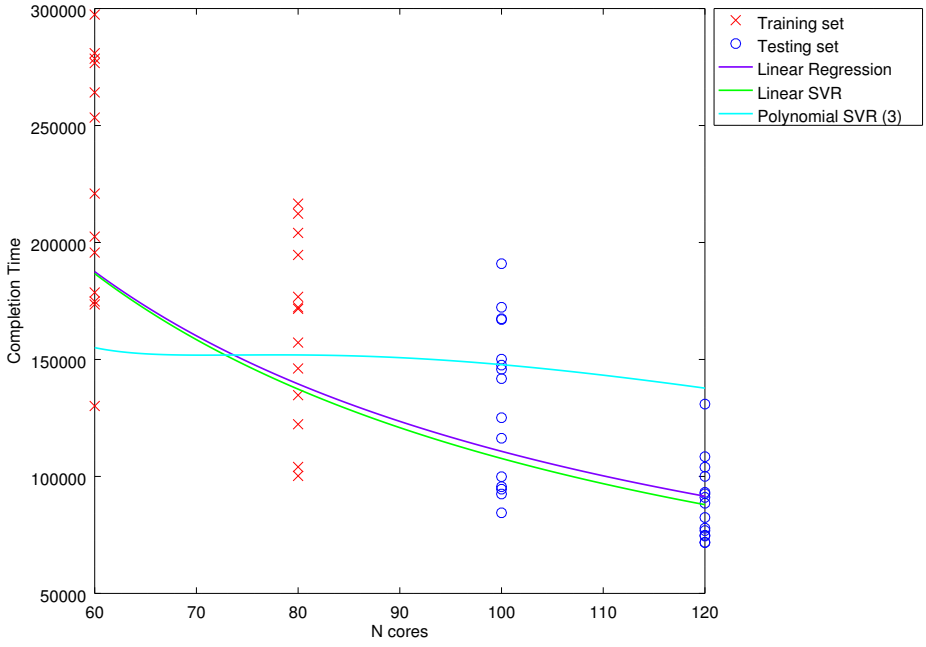

Figure 39: Completion time vs ncores for R4 (250GB), testing on 100, 120 cores

1.4.4 Query R4, Datasize 500GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.7096	0.4833	73366	0.0947
Linear SVR	1.1590	0.9824	154330	0.2424
Polynomial SVR (2)	1.7159	0.8837	222494	0.3431
Polynomial SVR (3)	1.4843	0.7897	208562	0.3399
Polynomial SVR (4)	1.6233	0.5178	211027	0.3261
Polynomial SVR (6)	1.6387	0.2592	224943	0.3613
Gaussian SVR	1.5067	0.4405	199252	0.3114

Table 40: Results for R4 (500GB), testing on 60, 80 cores

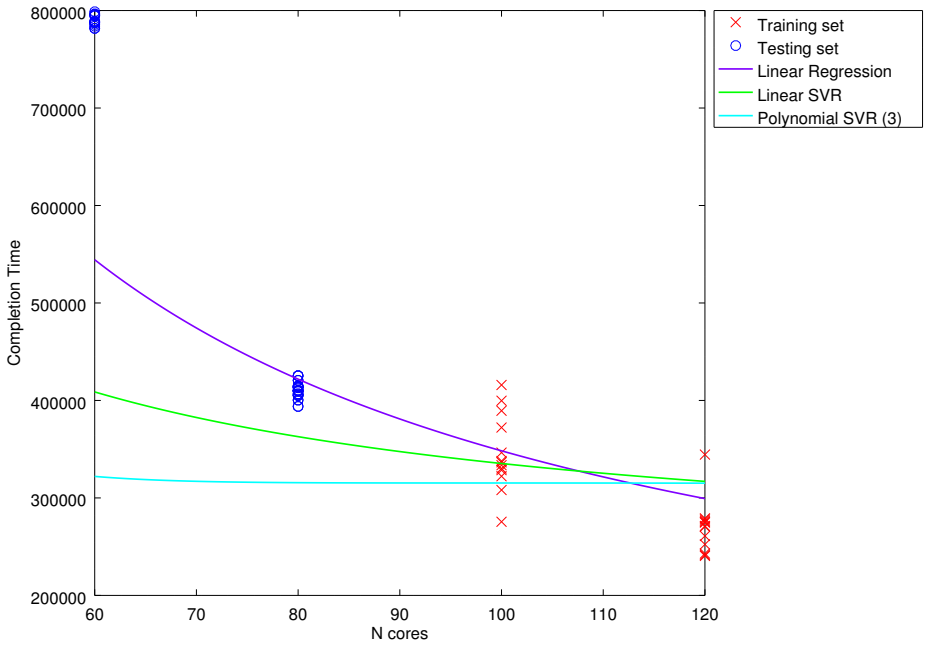

Figure 40: Completion time vs ncores for R4 (500GB), testing on 60, 80 cores

1.4.5 Query R4, Datasize 500GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5367	-5.2566	90415	0.2281
Linear SVR	0.2875	0.9428	47500	0.1195
Polynomial SVR (2)	0.9668	0.0207	161041	0.4069
Polynomial SVR (3)	0.1493	0.6233	24470	0.0652
Polynomial SVR (4)	0.5208	0.1051	85841	0.2145
Polynomial SVR (6)	0.5296	0.1980	87982	0.2192
Gaussian SVR	0.2911	0.6047	39492	0.1014

Table 41: Results for R4 (500GB), testing on 80, 100 cores

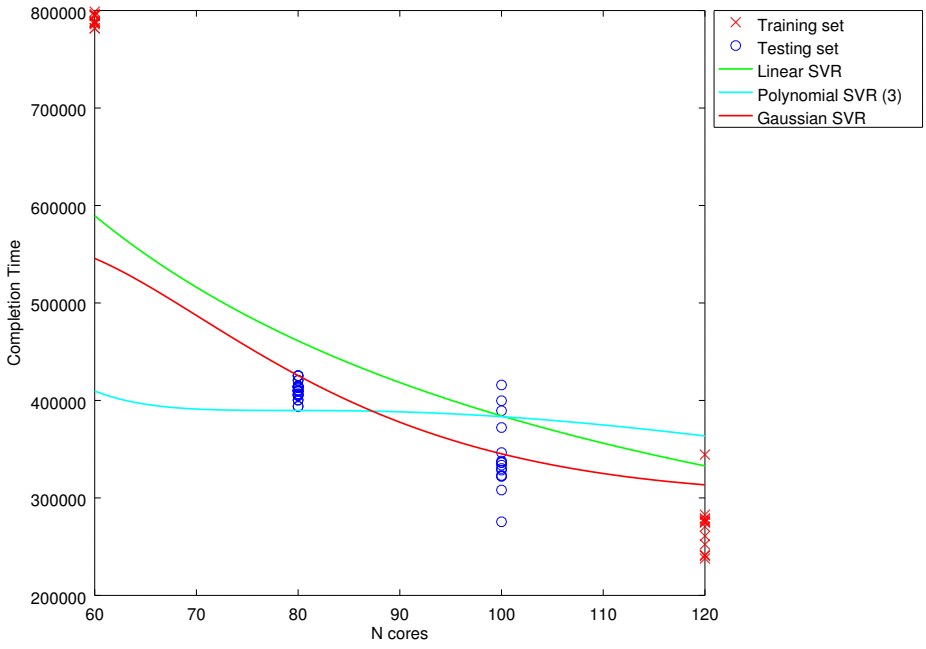

Figure 41: Completion time vs ncores for R4 (500GB), testing on 80, 100 cores

1.4.6 Query R4, Datasize 500GB — testing on 100, 120 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.7423	-7.1184	128480	0.4433
Linear SVR	0.2009	0.7700	29591	0.0986
Polynomial SVR (2)	1.2202	0.6590	197870	0.7065
Polynomial SVR (3)	0.4150	0.6359	70237	0.2436
Polynomial SVR (4)	0.7441	0.0299	120508	0.4258
Polynomial SVR (6)	0.6692	0.0823	108524	0.3846
Gaussian SVR	1.0298	0.0120	168742	0.5937

Table 42: Results for R4 (500GB), testing on 100, 120 cores

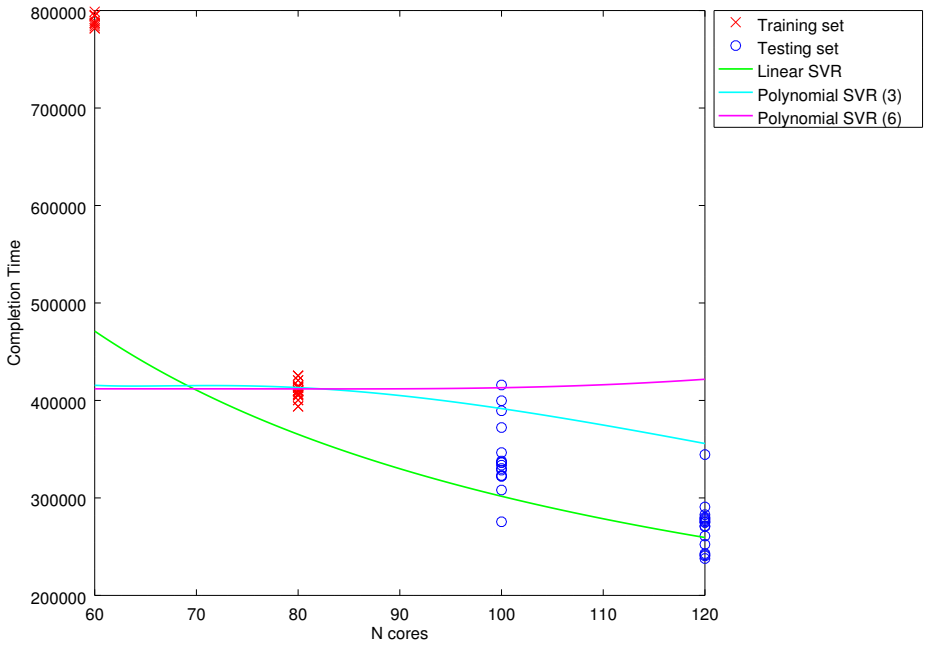

Figure 42: Completion time vs ncores for R4 (500GB), testing on 100, 120 cores

1.4.7 Query R4, Datasize 750GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.1791	-1.5620	142912	0.1906
Linear SVR	1.8072	0.0215	217501	0.2892
Polynomial SVR (2)	1.7981	0.0972	216714	0.2884
Polynomial SVR (3)	1.8278	0.1860	220224	0.2930
Polynomial SVR (4)	1.7806	-0.0000	213858	0.2842
Polynomial SVR (6)	1.7806	-0.0000	213858	0.2842
Gaussian SVR	1.7888	0.1385	215265	0.2862

Table 43: Results for R4 (750GB), testing on 60, 80 cores

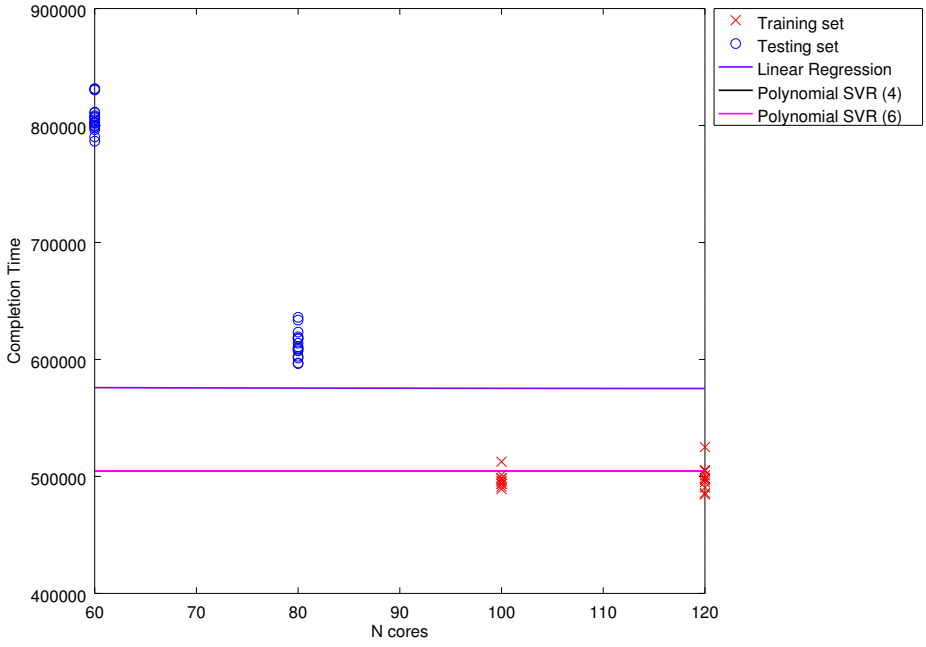

Figure 43: Completion time vs ncores for R4 (750GB), testing on 60, 80 cores

1.4.8 Query R4, Datasize 750GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.1961	0.8009	25055	0.0466
Linear SVR	0.2499	0.9841	30995	0.0580
Polynomial SVR (2)	1.9322	0.8239	212755	0.3600
Polynomial SVR (3)	0.3099	0.8259	33524	0.0645
Polynomial SVR (4)	2.9481	0.0074	152261	0.2588
Polynomial SVR (6)	6.7820	0.0026	220597	0.3878
Gaussian SVR	0.2565	0.8198	26332	0.0483

Table 44: Results for R4 (750GB), testing on 80, 100 cores

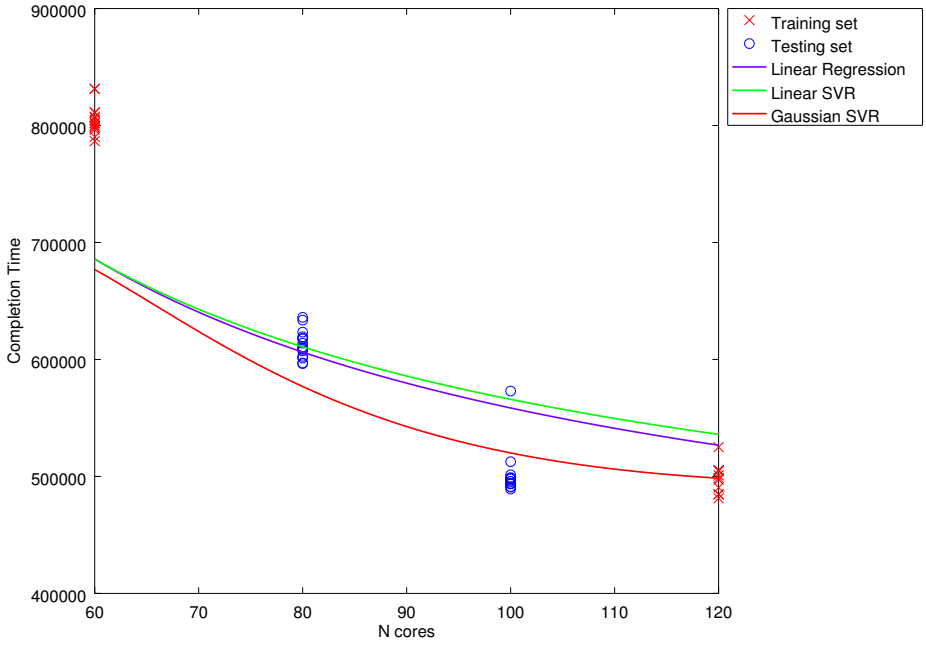

Figure 44: Completion time vs ncores for R4 (750GB), testing on 80, 100 cores

1.4.9 Query R4, Datasize 750GB — testing on 100, 120 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2318	-2.4894	22529	0.0453
Linear SVR	0.1325	0.6632	14405	0.0288
Polynomial SVR (2)	1.5032	0.2786	195044	0.3927
Polynomial SVR (3)	0.6336	0.4627	75255	0.1501
Polynomial SVR (4)	1.2186	0.1018	159225	0.3197
Polynomial SVR (6)	1.1516	0.4900	149151	0.2984
Gaussian SVR	1.3205	0.2683	171822	0.3444

Table 45: Results for R4 (750GB), testing on 100, 120 cores

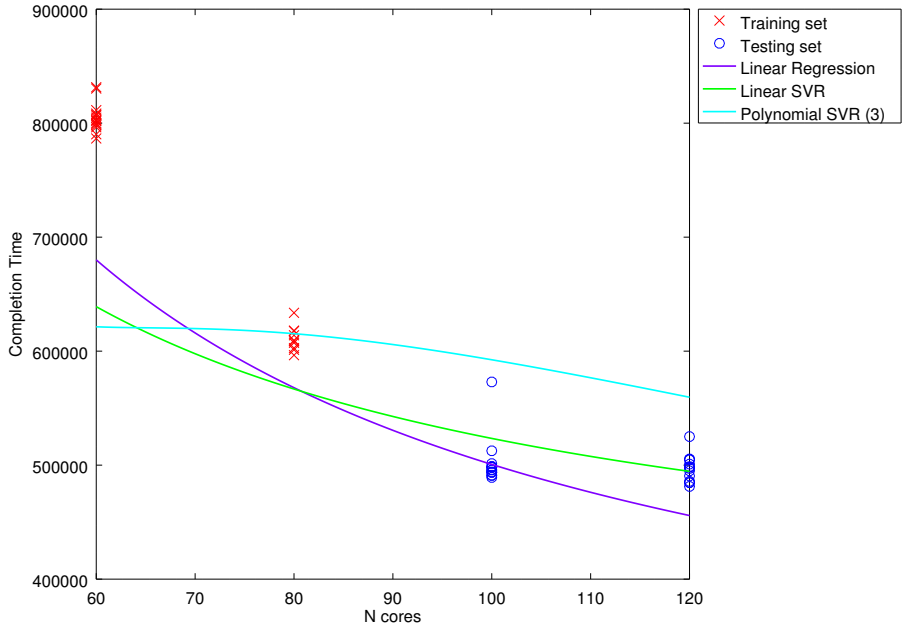

Figure 45: Completion time vs ncores for R4 (750GB), testing on 100, 120 cores

1.4.10 Query R4, Datasize 1000GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.5538	-4.5165	426665	0.3660
Linear SVR	0.2950	0.0886	176382	0.1384
Polynomial SVR (2)	1.7995	0.5068	1497888	1.4104
Polynomial SVR (3)	1.4826	0.0312	1241374	1.1663
Polynomial SVR (4)	1.7009	0.0120	1427968	1.3363
Polynomial SVR (6)	1.6854	0.0083	1413778	1.3246
Gaussian SVR	1.6227	0.5165	1368444	1.2701

Table 46: Results for R4 (1000GB), testing on 60, 80 cores

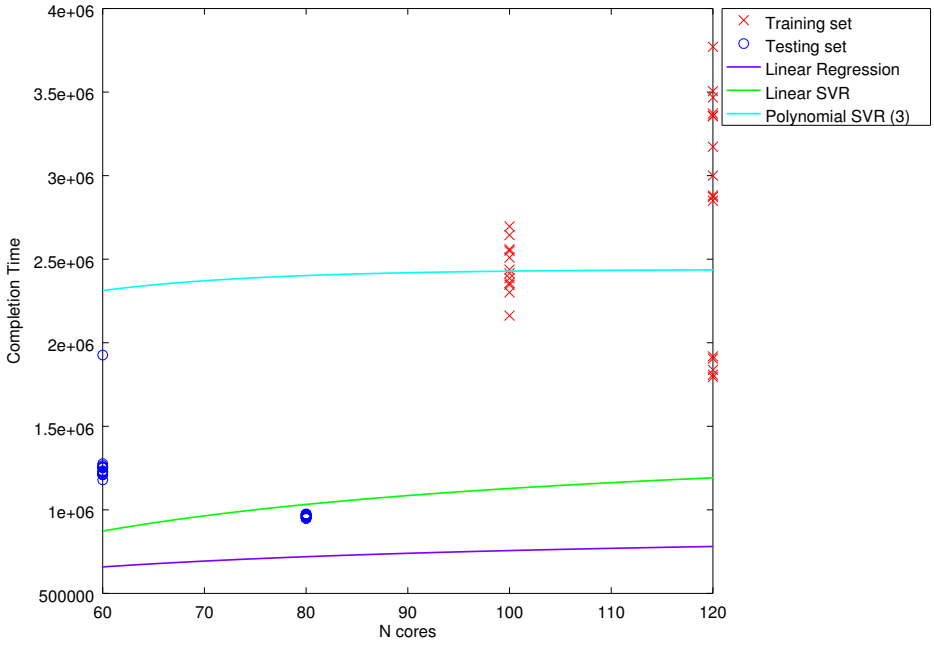

Figure 46: Completion time vs ncores for R4 (1000GB), testing on 60, 80 cores

1.4.11 Query R4, Datasize 1000GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2507	0.9159	139012	0.0723
Linear SVR	0.3655	0.8991	198491	0.1408
Polynomial SVR (2)	1.4002	0.3843	918738	0.8916
Polynomial SVR (3)	0.7984	0.4788	606892	0.5182
Polynomial SVR (4)	1.1687	0.1284	777513	0.7469
Polynomial SVR (6)	1.0167	0.0279	739083	0.6649
Gaussian SVR	0.6083	0.8712	411513	0.3838

Table 47: Results for R4 (1000GB), testing on 80, 100 cores

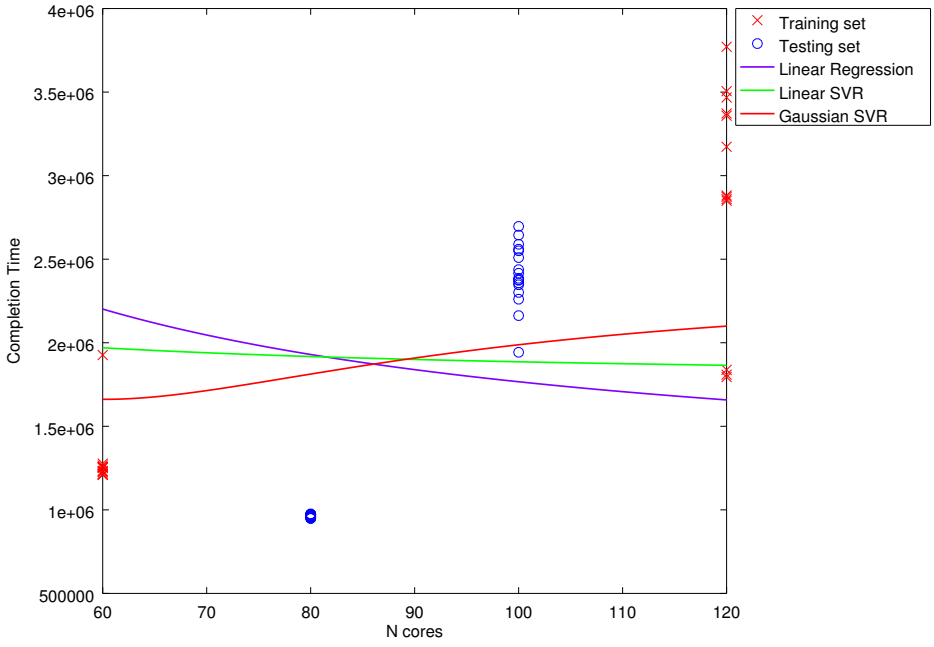

Figure 47: Completion time vs ncores for R4 (1000GB), testing on 80, 100 cores

1.4.12 Query R4, Datasize 1000GB — testing on 100, 120 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.4508	-4.2228	1175606	0.4452
Linear SVR	1.2732	0.1735	1013092	0.3922
Polynomial SVR (2)	1.9260	0.5189	1472033	0.5375
Polynomial SVR (3)	1.8734	0.4695	1516474	0.5749
Polynomial SVR (4)	1.9310	0.4082	1521459	0.5657
Polynomial SVR (6)	1.9556	0.4138	1543950	0.5751
Gaussian SVR	1.7530	0.2808	1392179	0.5201

Table 48: Results for R4 (1000GB), testing on 100, 120 cores

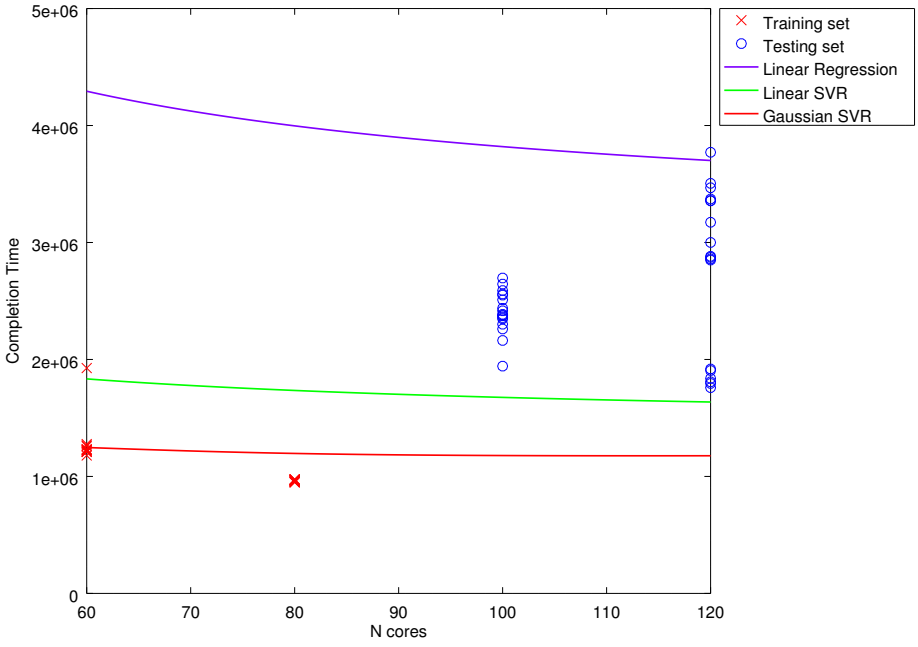

Figure 48: Completion time vs ncores for R4 (1000GB), testing on 100, 120 cores

1.5 Query R5

1.5.1 Query R5, Datasize 250GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.0902	0.0850	635	0.0246
Linear SVR	1.0697	0.2594	588	0.0229
Polynomial SVR (2)	1.2737	0.0284	703	0.0270
Polynomial SVR (3)	1.2741	0.0224	705	0.0271
Polynomial SVR (4)	1.2715	0.0170	702	0.0270
Polynomial SVR (6)	1.2709	0.0065	701	0.0270
Gaussian SVR	1.1492	0.1821	635	0.0245

Table 49: Results for R5 (250GB), testing on 60, 80 cores

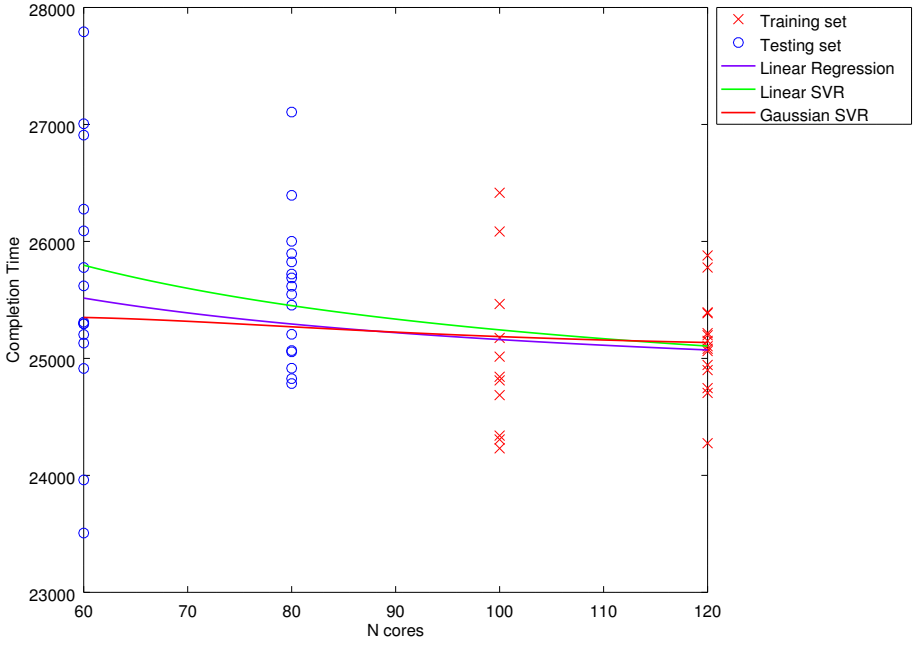

Figure 49: Completion time vs ncores for R5 (250GB), testing on 60, 80 cores

1.5.2 Query R5, Datasize 250GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.5079	-1.8906	713	0.0280
Linear SVR	0.9809	0.1001	642	0.0254
Polynomial SVR (2)	1.1878	0.0447	743	0.0297
Polynomial SVR (3)	1.0758	0.0233	663	0.0265
Polynomial SVR (4)	1.1439	0.0078	732	0.0292
Polynomial SVR (6)	1.1549	0.0007	731	0.0292
Gaussian SVR	1.0751	0.0121	669	0.0267

Table 50: Results for R5 (250GB), testing on 80, 100 cores

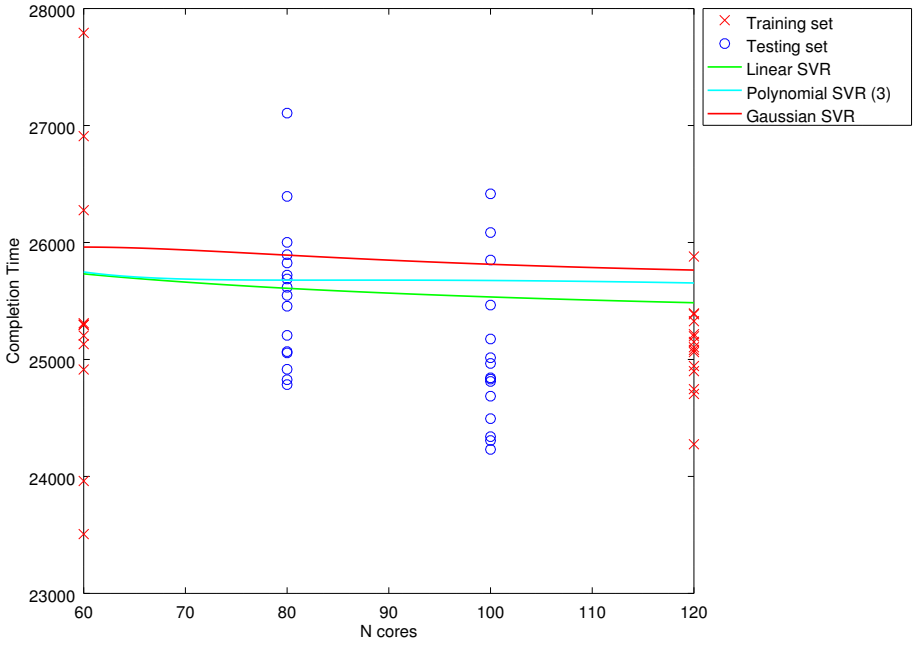

Figure 50: Completion time vs ncores for R5 (250GB), testing on 80, 100 cores

1.5.3 Query R5, Datasize 250GB — testing on 100, 120 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	1.3383	-2.7436	639	0.0253
Linear SVR	0.7851	0.0305	500	0.0199
Polynomial SVR (2)	1.0935	0.0233	688	0.0277
Polynomial SVR (3)	0.9228	0.0003	596	0.0239
Polynomial SVR (4)	1.0906	0.0153	709	0.0284
Polynomial SVR (6)	1.2867	0.0164	779	0.0312
Gaussian SVR	0.8820	0.1663	567	0.0228

Table 51: Results for R5 (250GB), testing on 100, 120 cores

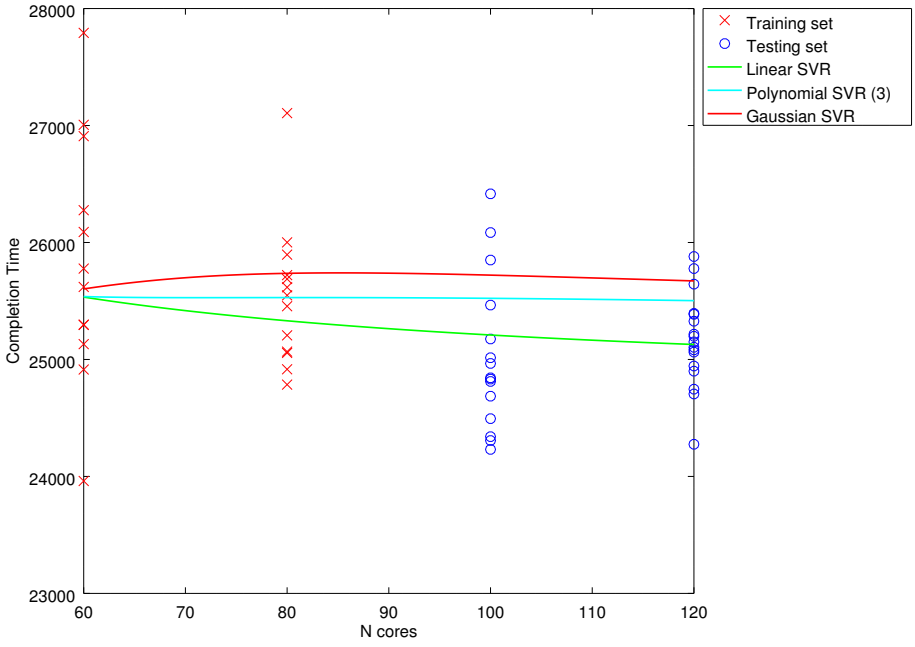

Figure 51: Completion time vs ncores for R5 (250GB), testing on 100, 120 cores

1.5.4 Query R5, Datasize 500GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4853	0.6844	438	0.0191
Linear SVR	0.4844	0.8048	427	0.0186
Polynomial SVR (2)	1.4614	0.1439	1389	0.0609
Polynomial SVR (3)	0.6057	0.5470	546	0.0237
Polynomial SVR (4)	2.1939	0.0847	1431	0.0631
Polynomial SVR (6)	1.7608	0.0552	1309	0.0575
Gaussian SVR	0.7085	0.6749	675	0.0298

Table 52: Results for R5 (500GB), testing on 60, 80 cores

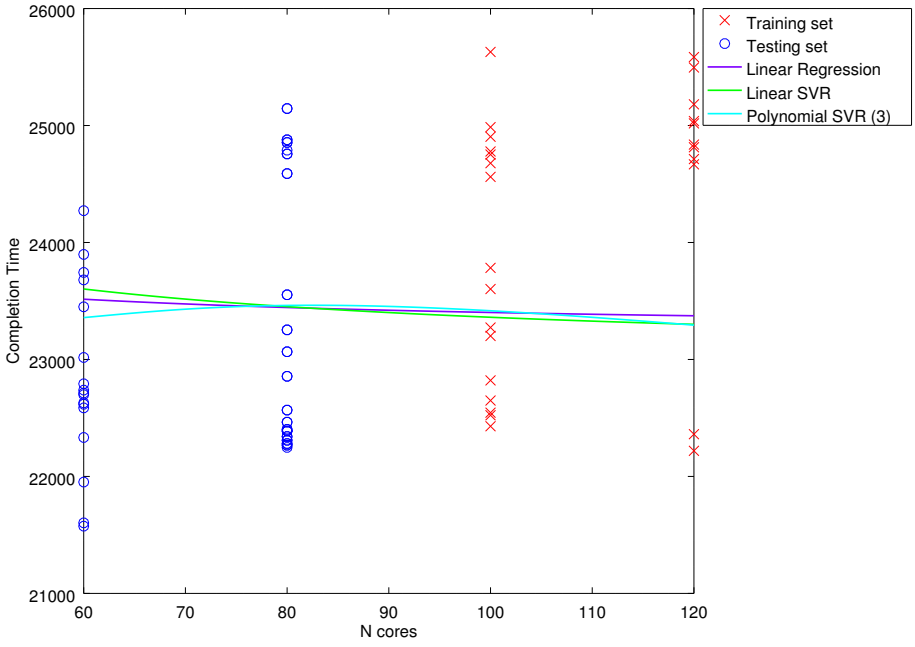

Figure 52: Completion time vs ncores for R5 (500GB), testing on 60, 80 cores

1.5.5 Query R5, Datasize 500GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.4098	0.8097	398	0.0170
Linear SVR	0.3012	0.8979	300	0.0129
Polynomial SVR (2)	0.8844	0.1690	927	0.0400
Polynomial SVR (3)	0.5981	0.7517	619	0.0265
Polynomial SVR (4)	0.8281	0.2351	861	0.0368
Polynomial SVR (6)	0.8828	0.1352	893	0.0377
Gaussian SVR	0.3482	0.8727	316	0.0134

Table 53: Results for R5 (500GB), testing on 80, 100 cores

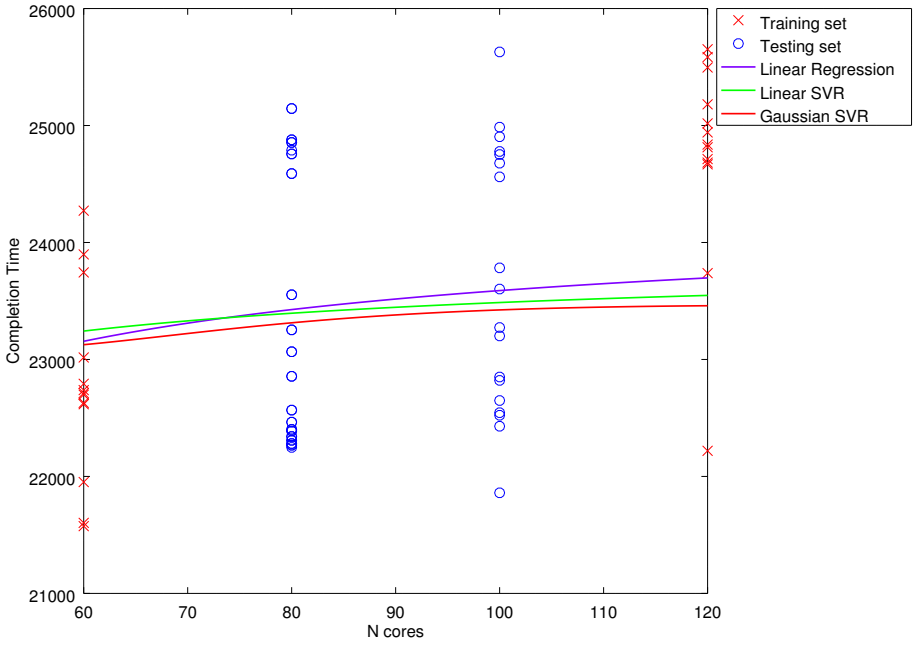

Figure 53: Completion time vs ncores for R5 (500GB), testing on 80, 100 cores

1.5.6 Query R5, Datasize 500GB — testing on 100, 120 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.2434	0.9423	240	0.0101
Linear SVR	0.3119	0.9366	300	0.0126
Polynomial SVR (2)	1.7141	0.5074	1731	0.0707
Polynomial SVR (3)	0.5719	0.6989	577	0.0240
Polynomial SVR (4)	1.5470	0.2546	1528	0.0620
Polynomial SVR (6)	1.1651	0.1858	1231	0.0507
Gaussian SVR	0.6641	0.7993	607	0.0248

Table 54: Results for R5 (500GB), testing on 100, 120 cores

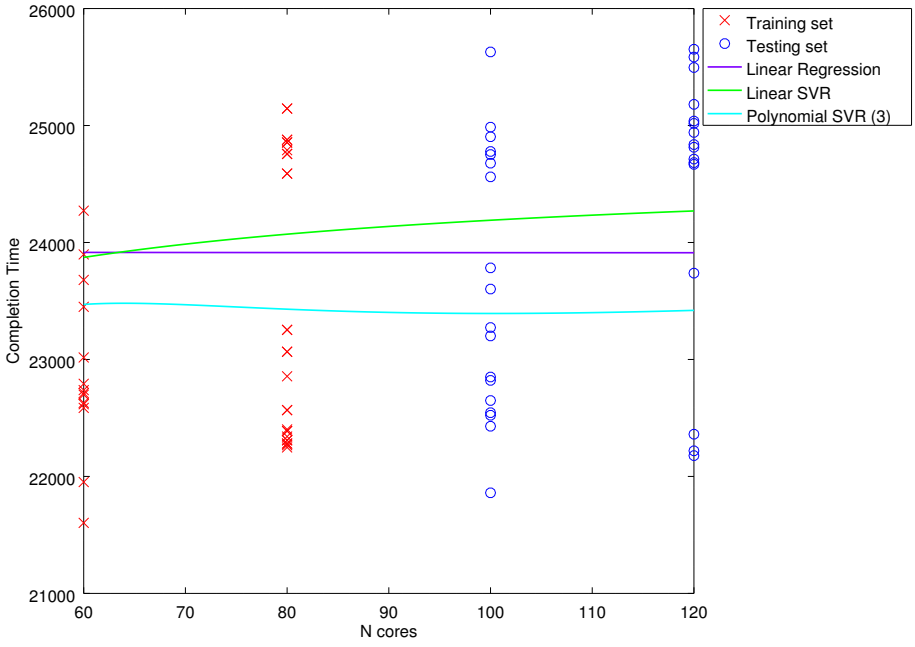

Figure 54: Completion time vs ncores for R5 (500GB), testing on 100, 120 cores

1.5.7 Query R5, Datasize 750GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.7333	0.2860	621	0.0259
Linear SVR	0.5451	0.6249	403	0.0169
Polynomial SVR (2)	1.3937	0.3825	1525	0.0649
Polynomial SVR (3)	0.5726	0.6847	568	0.0237
Polynomial SVR (4)	1.2149	0.3152	1323	0.0563
Polynomial SVR (6)	1.0959	0.0515	1202	0.0510
Gaussian SVR	0.6606	0.4645	595	0.0253

Table 55: Results for R5 (750GB), testing on 60, 80 cores

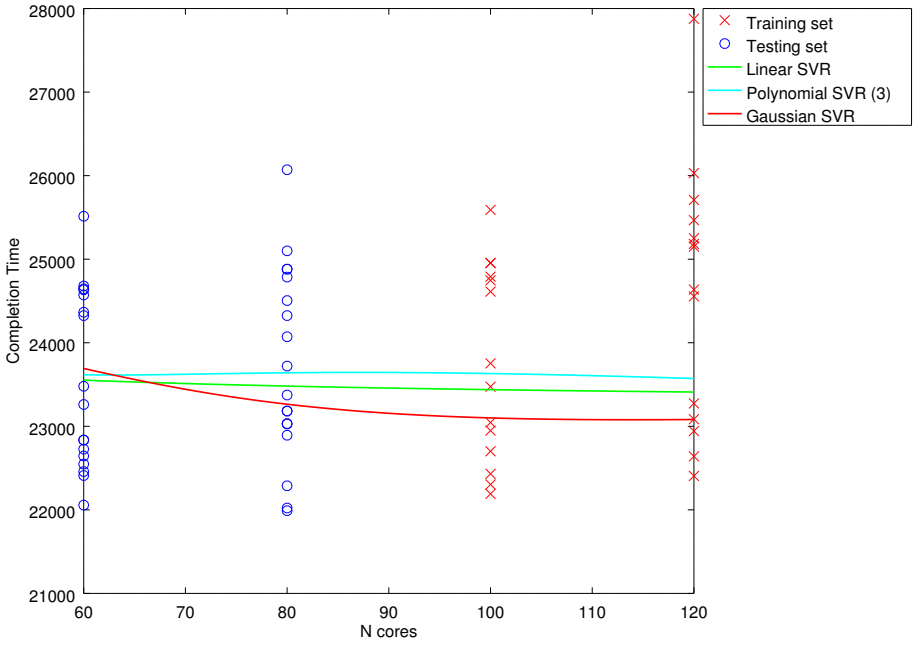

Figure 55: Completion time vs ncores for R5 (750GB), testing on 60, 80 cores

1.5.8 Query R5, Datasize 750GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.6448	0.4780	500	0.0207
Linear SVR	0.5907	0.6305	431	0.0178
Polynomial SVR (2)	1.0727	0.1608	1167	0.0497
Polynomial SVR (3)	0.7387	0.5108	780	0.0327
Polynomial SVR (4)	0.9341	0.1285	1035	0.0436
Polynomial SVR (6)	0.9175	0.0771	1023	0.0435
Gaussian SVR	0.6087	0.5363	495	0.0208

Table 56: Results for R5 (750GB), testing on 80, 100 cores

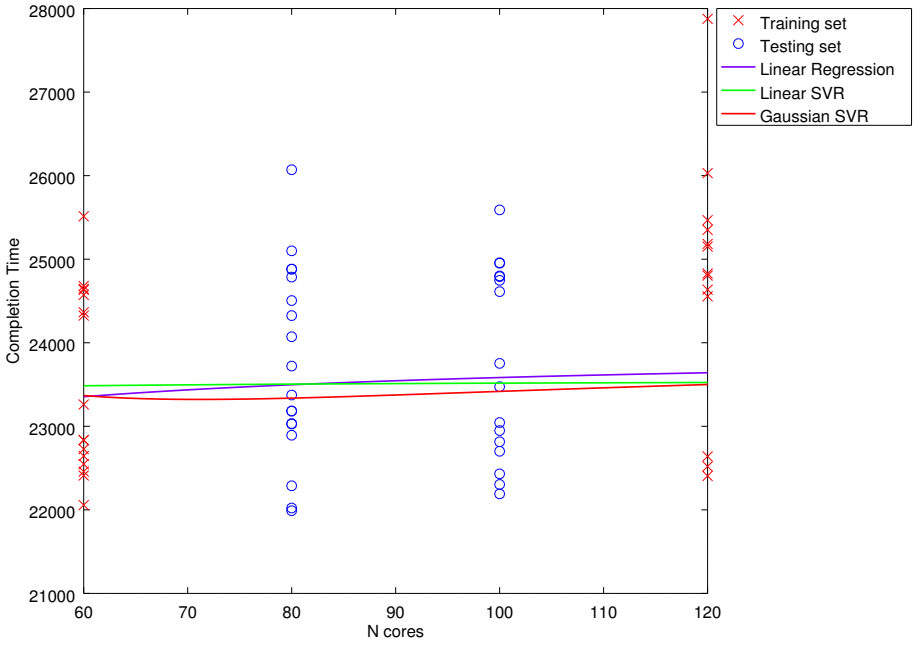

Figure 56: Completion time vs ncores for R5 (750GB), testing on 80, 100 cores

1.5.9 Query R5, Datasize 750GB — testing on 100, 120 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.9062	0.2779	781	0.0323
Linear SVR	0.6910	0.6206	529	0.0215
Polynomial SVR (2)	1.7067	0.2941	1836	0.0747
Polynomial SVR (3)	0.9975	0.3559	785	0.0321
Polynomial SVR (4)	1.6760	0.2247	1668	0.0674
Polynomial SVR (6)	2.3635	0.1594	1997	0.0804
Gaussian SVR	0.8882	0.3832	818	0.0335

Table 57: Results for R5 (750GB), testing on 100, 120 cores

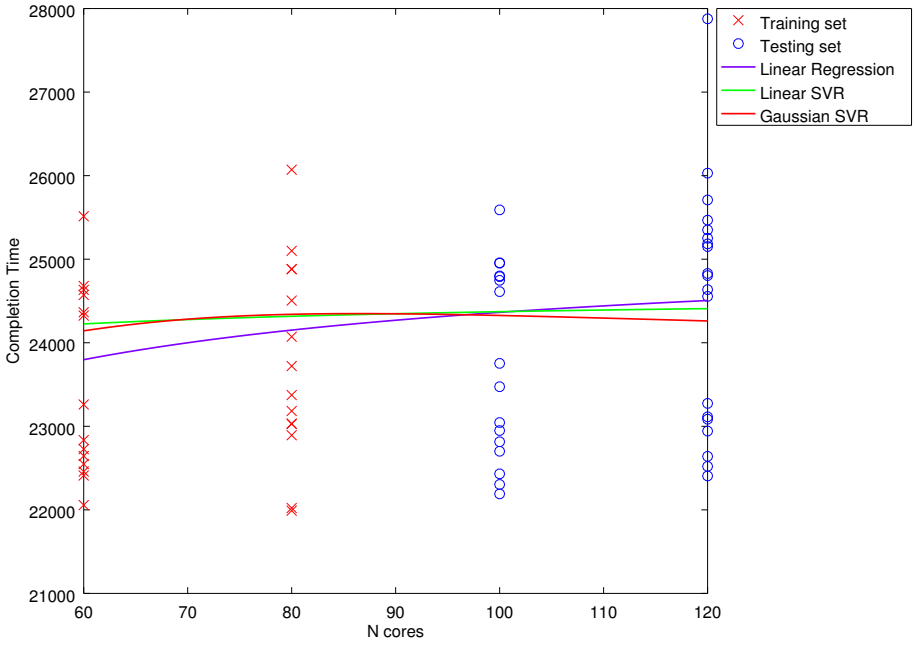

Figure 57: Completion time vs ncores for R5 (750GB), testing on 100, 120 cores

1.5.10 Query R5, Datasize 1000GB — testing on 60, 80 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	8.7159	-45.7583	24894	0.4599
Linear SVR	5.2648	0.9620	21496	0.4463
Polynomial SVR (2)	1.5531	0.7570	4398	0.0838
Polynomial SVR (3)	11.3789	0.6511	26536	0.4732
Polynomial SVR (4)	40.6531	0.5900	73576	1.2266
Polynomial SVR (6)	310.3589	0.5291	520118	8.5388
Gaussian SVR	1.5182	0.5927	6047	0.1244

Table 58: Results for R5 (1000GB), testing on 60, 80 cores

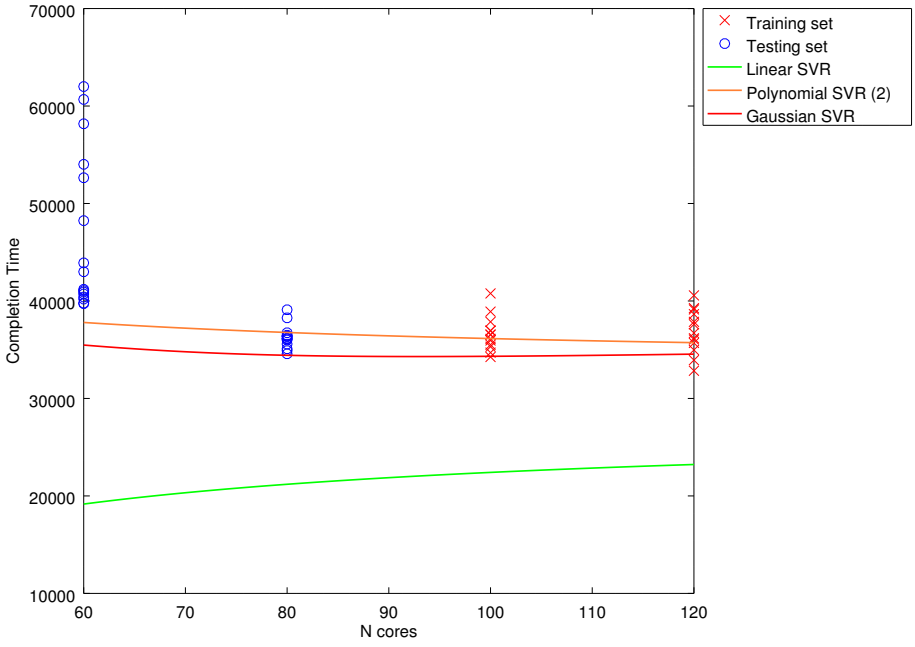

Figure 58: Completion time vs ncores for R5 (1000GB), testing on 60, 80 cores

1.5.11 Query R5, Datasize 1000GB — testing on 80, 100 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.3559	-1.1212	1892	0.0522
Linear SVR	0.3287	0.0758	1760	0.0485
Polynomial SVR (2)	0.4317	0.1574	2237	0.0601
Polynomial SVR (3)	0.3475	0.2887	1881	0.0518
Polynomial SVR (4)	0.2423	0.1934	1051	0.0281
Polynomial SVR (6)	0.2512	0.0995	1267	0.0346
Gaussian SVR	0.2168	0.4127	1097	0.0300

Table 59: Results for R5 (1000GB), testing on 80, 100 cores

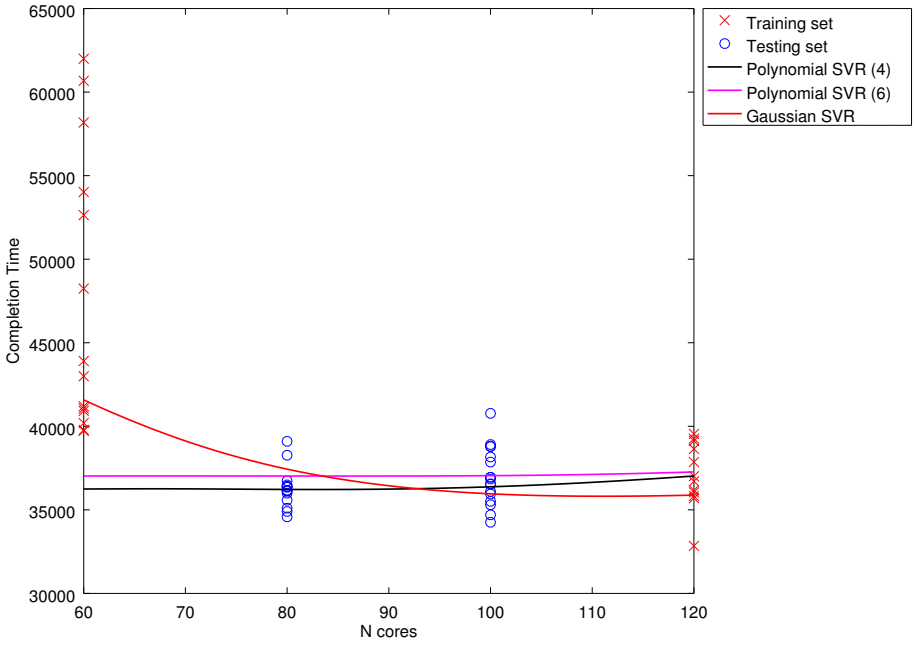

Figure 59: Completion time vs ncores for R5 (1000GB), testing on 80, 100 cores

1.5.12 Query R5, Datasize 1000GB — testing on 100, 120 cores

Model	RMSE	R^2	Mean absolute error	Mean relative error
Linear regression	0.7663	-5.1490	3926	0.1032
Linear SVR	0.7032	0.4710	3593	0.0944
Polynomial SVR (2)	0.3831	0.4884	1995	0.0554
Polynomial SVR (3)	0.4646	0.3326	2193	0.0578
Polynomial SVR (4)	0.2777	0.2829	1345	0.0367
Polynomial SVR (6)	0.3037	0.2164	1523	0.0417
Gaussian SVR	0.4800	0.2897	2528	0.0701

Table 60: Results for R5 (1000GB), testing on 100, 120 cores

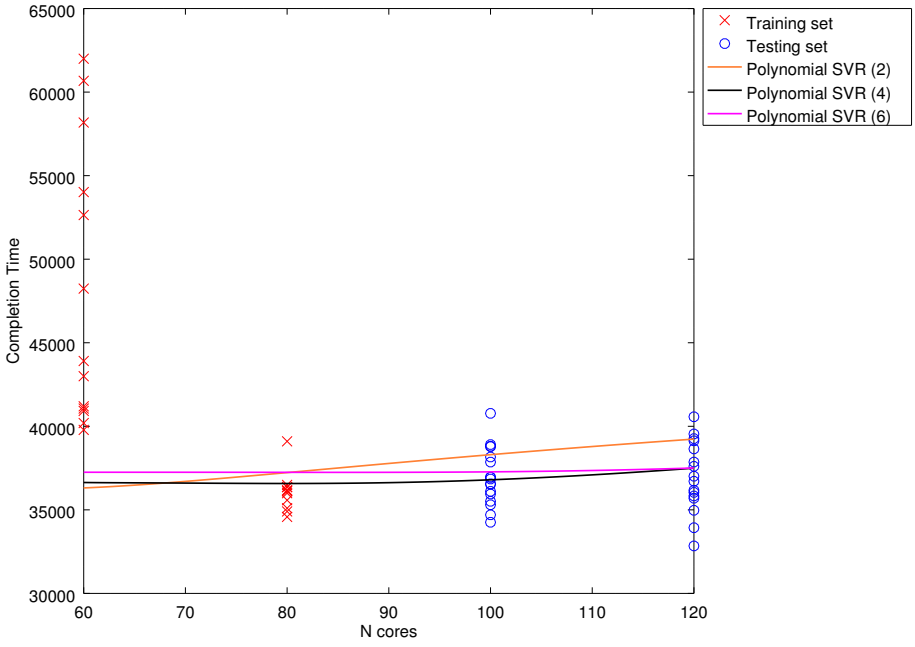

Figure 60: Completion time vs ncores for R5 (1000GB), testing on 100, 120 cores