

Estimation of parameter importance and interaction with Random Forests

Leslie Pérez Cáceres

In this document we are testing how to use Random Forest to assess importance and interactions of parameters based on the data gathered by irace.

For the analysis below we use as example ACOTSP in the short scenario:

- 7 secs cut off time
- 11 parameters
- 400 instances sizes (1000,1500,2000,2500,3000,3500,4000,4500)
- 3000 experiments configuration budget

1. Importance by iterations

Iteration 1:

	variable	mean_min_depth	no_of_nodes	mse_increase	node_purity_increase	no_of_trees	times_a_root	p_value
9	localsearch	1.440000	1209	516.48990	145408.893	300	94	1.0000000
4	beta	2.156667	4500	82.52202	25000.856	300	30	0.0000000
2	alpha	2.360000	4417	69.66845	21281.047	300	30	0.0000000
8	instance	2.396667	10035	80.53465	39281.125	300	0	0.0000000
10	nnls	2.406667	2976	267.80183	72673.210	300	60	0.9999856
1	algorithm	2.916667	2754	47.34339	11443.078	300	1	1.0000000
13	rho	2.943333	3917	38.11551	11070.242	300	4	0.0000000
3	ants	3.036667	3904	30.47331	9190.517	300	3	0.0000000
6	dummy	3.360000	3401	24.28858	7095.121	300	6	0.0001325
5	dlb	3.644400	1046	206.34644	55339.015	297	50	1.0000000
12	rasrank	3.883333	1305	21.81376	6069.397	300	21	1.0000000
7	elitistants	4.088044	897	24.81032	6130.291	293	0	1.0000000
11	q0	4.441822	1246	12.56371	4072.382	298	1	1.0000000

Iteration 2:

	variable	mean_min_depth	no_of_nodes	mse_increase	node_purity_increase	no_of_trees	times_a_root	p_value
9	localsearch	1.490000	1118	327.63583	95805.624	300	90	1.0000000
8	instance	2.110000	10836	144.60159	60630.166	300	0	0.0000000
10	nnls	2.403333	3678	154.11492	43730.788	300	60	0.0000229
1	algorithm	2.543333	2088	77.54690	22058.661	300	21	1.0000000
7	elitistants	2.797422	1416	58.82252	16572.166	298	10	1.0000000
2	alpha	2.840000	4538	26.40868	11337.389	300	14	0.0000000
13	rho	2.923333	4658	47.53169	14502.564	300	0	0.0000000
6	dummy	3.006667	3416	27.44354	9475.040	300	5	0.7017641
11	q0	3.087422	1310	45.93340	14152.511	298	21	1.0000000
5	dlb	3.114467	1203	139.82902	39096.800	297	71	1.0000000
3	ants	3.130000	4615	26.85838	10562.283	300	1	0.0000000
4	beta	3.260000	4499	12.01566	5672.181	300	0	0.0000000
12	rasrank	4.313711	1416	17.88547	7186.986	299	7	1.0000000

Iteration 3:

	variable	mean_min_depth	no_of_nodes	mse_increase	node_purity_increase	no_of_trees	times_a_root	p_value
9	localsearch	1.642722	712	192.56629	46116.849	299	91	1.0000000
8	instance	2.213333	10768	41.83292	35954.947	300	0	0.0000000
10	nnls	2.226667	3764	100.20869	27692.614	300	65	0.0000000
3	ants	2.696667	4070	95.20598	23758.917	300	24	0.0000000
2	alpha	2.740000	4168	98.43716	25487.817	300	1	0.0000000
12	rasrank	2.833333	1883	87.28371	21544.918	300	30	1.0000000
4	beta	2.853333	4024	69.00544	19130.870	300	9	0.0000000
1	algorithm	2.980000	1726	106.99919	19521.411	300	20	1.0000000
13	rho	3.340000	4063	57.86721	16352.399	300	0	0.0000000
6	dummy	3.433333	3189	40.28508	9456.497	300	3	0.3096768
5	dlb	3.582111	1155	49.70874	12432.818	298	38	1.0000000
7	elitistants	3.953389	518	30.69426	7822.016	251	19	1.0000000
11	q0	5.095722	1063	26.01928	5923.417	293	0	1.0000000

Iteration 4:

	variable	mean_min_depth	no_of_nodes	mse_increase	node_purity_increase	no_of_trees	times_a_root	p_value
1	algorithm	1.710000	1761	239.20049	56427.026	300	84	1.0000000
10	nnls	2.153333	4154	169.63543	41608.594	300	34	0.0000000
9	localsearch	2.221367	680	108.00097	30706.271	293	77	1.0000000
2	alpha	2.233333	4624	161.78526	49609.314	300	33	0.0000000
8	instance	2.403333	12524	24.53823	46459.338	300	0	0.0000000
3	ants	2.496667	4276	130.83339	35920.178	300	16	0.0000000
13	rho	2.583333	4329	110.73182	32691.796	300	11	0.0000000
4	beta	3.023333	4309	83.51097	26054.540	300	2	0.0000000
6	dummy	3.146667	3401	57.44221	17549.046	300	3	0.1402052
7	elitistants	4.202900	430	34.67711	9173.709	251	37	1.0000000
11	q0	4.753333	1973	48.45441	14488.980	300	0	1.0000000
5	dlb	6.209833	401	13.16681	4987.288	235	3	1.0000000
12	rasrank	6.654633	565	10.77075	3927.525	247	0	1.0000000

Iteration 5:

	variable	mean_min_depth	no_of_nodes	mse_increase	node_purity_increase	no_of_trees	times_a_root	p_value
3	ants	1.263333	7235	571.992805	198817.125	300	94	0
10	nnls	1.986667	7450	262.650023	127546.677	300	39	0
2	alpha	1.996667	7977	246.282891	117999.329	300	53	0
12	rho	2.136667	6850	230.143899	89640.707	300	48	0
8	instance	2.410000	19813	77.931583	170096.430	300	0	0
11	rasrank	2.713333	1969	146.216577	61612.332	300	7	1
1	algorithm	2.993833	886	200.208017	46707.292	285	40	1
4	beta	2.996667	7776	126.438934	75593.318	300	0	0
6	dummy	3.096667	5925	86.983774	43627.721	300	0	0
9	localsearch	5.176833	248	7.402569	5915.474	195	11	1
5	dlb	5.951411	183	6.117048	4406.353	151	8	1
7	elitistants	6.489800	69	3.333008	2609.622	66	0	1

Iteration 6:

	variable	mean_min_depth	no_of_nodes	mse_increase	node_purity_increase	no_of_trees	times_a_root	p_value
3	ants	1.910000	3085	24.954673	5417.3251	300	60	0.8931291
6	instance	1.936667	9207	5.221205	14049.5960	300	2	0.0000000
7	nnls	2.114378	2668	23.592434	3871.4933	299	59	1.0000000
2	alpha	2.463333	3332	20.609703	4744.7731	300	43	0.0003368
8	rasrank	2.476711	704	13.254185	3017.5817	284	67	1.0000000
4	beta	2.696667	3279	15.411162	4054.4772	300	16	0.0079235
9	rho	2.710000	2945	15.835968	3881.4797	300	37	0.9999563
5	dummy	2.940000	2822	12.761248	3572.1667	300	16	1.0000000
1	algorithm	6.678622	310	1.768370	403.9089	196	0	1.0000000

Iteration 7:

	variable	mean_min_depth	no_of_nodes	mse_increase	node_purity_increase	no_of_trees	times_a_root	p_value
6	instance	1.821549	1736	-0.4973019	108.288593	297	6	0.0000000
3	ants	2.514950	511	0.6048508	46.123911	255	89	0.5850719
9	rho	2.698709	444	0.5376478	46.039061	242	90	0.9996843
7	nnls	2.890269	454	0.4441508	40.242552	234	71	0.9983115
5	dummy	3.442638	424	0.0235428	9.273771	238	12	0.9999946
2	alpha	3.706319	417	-0.0139712	10.746582	224	3	0.9999990
4	beta	3.731661	431	0.0044984	10.185637	226	14	0.9999749
8	rasrank	4.770045	118	0.0439301	6.175538	106	10	1.0000000
1	algorithm	4.804770	102	0.0318831	5.180587	92	5	1.0000000

Iteration 8:

	variable	mean_min_depth	no_of_nodes	mse_increase	node_purity_increase	no_of_trees	times_a_root	p_value
6	instance	1.118056	1412	-0.1374562	61.168989	288	85	0.0000000
2	alpha	2.366250	387	0.0940148	4.730175	227	60	0.4744831
4	beta	2.838472	340	0.0694194	3.407910	202	31	0.9941407
9	rho	3.058889	302	0.0550899	2.819211	191	29	0.9999986
7	nnls	3.082222	288	0.0574791	2.993466	187	23	1.0000000
3	ants	3.260833	270	0.0366201	2.437394	179	21	1.0000000
5	dummy	3.313889	266	0.0157581	2.131594	183	6	1.0000000
8	rasrank	3.722917	102	0.0361496	2.076050	98	23	1.0000000
1	algorithm	3.760417	102	0.0437524	2.273046	88	22	1.0000000

The 5 most important parameters per iteration are :

```
## [[1]]
## [1] "localsearch" "beta"          "alpha"          "instance"
##
## [[2]]
## [1] "localsearch" "instance"      "algorithm"      "alpha"
##
## [[3]]
## [1] "localsearch" "instance"      "ants"          "alpha"
##
## [[4]]
## [1] "algorithm"   "nnls"        "localsearch" "alpha"        "instance"
##
## [[5]]
## [1] "ants"        "nnls"        "alpha"        "rho"          "instance"
##
## [[6]]
## [1] "ants"        "instance"    "nnls"        "alpha"        "beta"
##
## [[7]]
## [1] "instance"    "ants"        "rho"         "nnls"
##
## [[8]]
## [1] "instance"    "alpha"       "beta"        "rho"          "nnls"
```

Parameter relevant interactions:

Iteration 1:

variable	root_variable	mean_min_depth	occurrences	interaction	uncond_mean_min_depth
beta	localsearch	0.9963768	516	localsearch:beta	1.513333

Iteration 2:

variable	root_variable	mean_min_depth	occurrences	interaction	uncond_mean_min_depth
----------	---------------	----------------	-------------	-------------	-----------------------

Iteration 3:

variable	root_variable	mean_min_depth	occurrences	interaction	uncond_mean_min_depth
alpha	ants	1.114082	512	ants:alpha	1.446667

Iteration 4:

variable	root_variable	mean_min_depth	occurrences	interaction	uncond_mean_min_depth
nnls	localsearch	1.073571	450	localsearch:nnls	1.553333
alpha	localsearch	1.288571	436	localsearch:alpha	1.703333

Iteration 5:

variable	root_variable	mean_min_depth	occurrences	interaction	uncond_mean_min_depth
ants	alpha	1.312369	552	alpha:ants	1.320000
rho	alpha	1.413415	545	alpha:rho	1.650000
nnls	alpha	1.485854	555	alpha:nnls	1.526667

Iteration 6:

variable	root_variable	mean_min_depth	occurrences	interaction	uncond_mean_min_depth
----------	---------------	----------------	-------------	-------------	-----------------------

Iteration 7:

variable	root_variable	mean_min_depth	occurrences	interaction	uncond_mean_min_depth
----------	---------------	----------------	-------------	-------------	-----------------------

Iteration 8:

variable	root_variable	mean_min_depth	occurrences	interaction	uncond_mean_min_depth
----------	---------------	----------------	-------------	-------------	-----------------------