Random Forest analysis of the cars dataset predicting the Make of the car

The HPFOREST Procedure

Performance Information	
Execution Mode	Single-Machine
Number of Threads	4

Data Access Information			
Data	**Engine**	**Role**	**Path**
WORK.CARS	V9	Input	On Client

Model Information		
Parameter	**Value**	
Variables to Try	4	(Default)
Maximum Trees	100	(Default)
Actual Trees	100	
Inbag Fraction	0.6	(Default)
Prune Fraction	0	(Default)
Prune Threshold	0.1	(Default)
Leaf Fraction	0.00001	(Default)
Leaf Size Setting	1	(Default)
Leaf Size Used	1	
Category Bins	30	(Default)
Interval Bins	100	
Minimum Category Size	5	(Default)
Node Size	100000	(Default)
Maximum Depth	20	(Default)
Alpha	1	(Default)
Exhaustive	5000	(Default)
Rows of Sequence to Skip	5	(Default)
Split Criterion	.	Gini
Preselection Method	.	Loh
Missing Value Handling	.	Valid value

Number of Observations	
Type	**N**
Number of Observations Read	428
Number of Observations Used	428

Random Forest analysis of the cars dataset predicting the Make of the car

The HPFOREST Procedure

Baseline Fit Statistics	
Statistic	**Value**
Average Square Error	0.025
Misclassification Rate	0.935
Log Loss	3.447

Random Forest analysis of the cars dataset predicting the Make of the car

The HPFOREST Procedure

Fit Statistics							
Number of Trees	Number of Leaves	Average Square Error (Train)	Average Square Error (OOB)	Misclassification Rate (Train)	Misclassification Rate (OOB)	Log Loss (Train)	Log Loss (OOB)
1	158	0.01442	0.0354	0.28037	0.680	6.254	15.54
2	297	0.00817	0.0308	0.27336	0.639	2.090	13.68
3	452	0.00609	0.0282	0.14486	0.604	0.963	12.56
4	602	0.00530	0.0263	0.08645	0.620	0.655	11.60
5	756	0.00475	0.0253	0.06308	0.617	0.400	10.77
6	901	0.00440	0.0236	0.05140	0.591	0.345	9.90
7	1047	0.00412	0.0230	0.02336	0.579	0.340	9.60
8	1197	0.00400	0.0220	0.03037	0.570	0.340	9.04
9	1348	0.00381	0.0212	0.01636	0.565	0.333	8.50
10	1494	0.00364	0.0204	0.02103	0.565	0.326	7.76
11	1638	0.00357	0.0198	0.01168	0.547	0.326	7.40
12	1785	0.00348	0.0188	0.01636	0.511	0.324	6.83
13	1947	0.00340	0.0185	0.00701	0.509	0.322	6.38
14	2093	0.00332	0.0183	0.00467	0.502	0.320	6.15
15	2248	0.00322	0.0179	0.00234	0.488	0.316	6.04
16	2400	0.00317	0.0176	0.00234	0.493	0.315	5.95
17	2554	0.00310	0.0174	0.00234	0.488	0.312	5.80
18	2703	0.00307	0.0173	0.00234	0.491	0.310	5.60
19	2861	0.00303	0.0171	0.00467	0.486	0.310	5.27
20	3015	0.00299	0.0170	0.00234	0.488	0.309	5.13
21	3182	0.00296	0.0170	0.00234	0.491	0.308	5.04
22	3332	0.00293	0.0169	0.00234	0.491	0.308	4.95
23	3489	0.00289	0.0168	0.00000	0.488	0.306	4.70
24	3650	0.00288	0.0168	0.00000	0.493	0.307	4.62
25	3810	0.00286	0.0167	0.00000	0.495	0.307	4.62
26	3960	0.00286	0.0166	0.00000	0.493	0.306	4.53
27	4123	0.00284	0.0166	0.00000	0.486	0.307	4.48
28	4281	0.00281	0.0164	0.00000	0.486	0.306	4.38
29	4439	0.00281	0.0164	0.00234	0.481	0.306	4.33
30	4579	0.00278	0.0163	0.00000	0.477	0.305	4.32
31	4728	0.00279	0.0163	0.00234	0.474	0.305	4.32
32	4872	0.00281	0.0164	0.00234	0.481	0.306	4.28
33	5020	0.00280	0.0163	0.00234	0.486	0.307	4.14

Random Forest analysis of the cars dataset predicting the Make of the car

The HPFOREST Procedure

Fit Statistics							
Number of Trees	Number of Leaves	Average Square Error (Train)	Average Square Error (OOB)	Misclassification Rate (Train)	Misclassification Rate (OOB)	Log Loss (Train)	Log Loss (OOB)
34	**5180**	0.00279	0.0162	0.00234	0.486	0.307	3.90
35	**5334**	0.00278	0.0162	0.00234	0.472	0.306	3.85
36	**5487**	0.00278	0.0161	0.00000	0.474	0.306	3.85
37	**5652**	0.00276	0.0161	0.00000	0.465	0.306	3.76
38	**5819**	0.00276	0.0161	0.00000	0.465	0.307	3.72
39	**5974**	0.00275	0.0161	0.00000	0.465	0.307	3.72
40	**6134**	0.00274	0.0161	0.00000	0.474	0.306	3.72
41	**6285**	0.00273	0.0160	0.00000	0.465	0.306	3.72
42	**6430**	0.00272	0.0160	0.00000	0.460	0.306	3.62
43	**6588**	0.00273	0.0161	0.00000	0.472	0.306	3.63
44	**6729**	0.00274	0.0160	0.00000	0.470	0.307	3.58
45	**6881**	0.00273	0.0160	0.00000	0.465	0.306	3.54
46	**7021**	0.00273	0.0160	0.00000	0.460	0.307	3.49
47	**7173**	0.00273	0.0160	0.00000	0.460	0.307	3.40
48	**7331**	0.00272	0.0159	0.00000	0.463	0.306	3.40
49	**7482**	0.00271	0.0159	0.00000	0.463	0.305	3.30
50	**7644**	0.00271	0.0159	0.00000	0.460	0.305	3.30
51	**7814**	0.00271	0.0159	0.00000	0.460	0.306	3.31
52	**7970**	0.00270	0.0159	0.00000	0.460	0.305	3.31
53	**8124**	0.00271	0.0159	0.00000	0.458	0.306	3.31
54	**8273**	0.00271	0.0159	0.00000	0.451	0.306	3.31
55	**8429**	0.00271	0.0159	0.00000	0.453	0.306	3.27
56	**8575**	0.00271	0.0158	0.00000	0.456	0.306	3.26
57	**8734**	0.00270	0.0158	0.00000	0.451	0.306	3.22
58	**8888**	0.00269	0.0157	0.00000	0.453	0.305	3.17
59	**9038**	0.00268	0.0157	0.00000	0.458	0.304	3.07
60	**9194**	0.00267	0.0156	0.00000	0.458	0.304	2.97
61	**9347**	0.00267	0.0156	0.00000	0.458	0.304	2.97
62	**9493**	0.00267	0.0156	0.00000	0.451	0.304	2.97
63	**9658**	0.00267	0.0156	0.00000	0.453	0.304	2.97
64	**9826**	0.00267	0.0156	0.00000	0.453	0.305	2.98
65	**9973**	0.00266	0.0155	0.00000	0.442	0.305	2.93
66	**10135**	0.00267	0.0156	0.00000	0.442	0.305	2.93

Random Forest analysis of the cars dataset predicting the Make of the car

The HPFOREST Procedure

Fit Statistics							
Number of Trees	Number of Leaves	Average Square Error (Train)	Average Square Error (OOB)	Misclassification Rate (Train)	Misclassification Rate (OOB)	Log Loss (Train)	Log Loss (OOB)
67	**10291**	0.00266	0.0155	0.00000	0.444	0.304	2.88
68	**10447**	0.00265	0.0155	0.00000	0.439	0.304	2.88
69	**10607**	0.00265	0.0155	0.00000	0.439	0.304	2.88
70	**10756**	0.00265	0.0155	0.00000	0.446	0.304	2.88
71	**10899**	0.00266	0.0155	0.00000	0.437	0.305	2.88
72	**11048**	0.00266	0.0155	0.00000	0.435	0.305	2.88
73	**11205**	0.00265	0.0155	0.00000	0.428	0.304	2.88
74	**11356**	0.00265	0.0155	0.00000	0.428	0.305	2.88
75	**11512**	0.00265	0.0155	0.00000	0.425	0.304	2.84
76	**11667**	0.00264	0.0155	0.00000	0.423	0.305	2.75
77	**11833**	0.00264	0.0155	0.00000	0.425	0.305	2.70
78	**11985**	0.00265	0.0155	0.00000	0.430	0.305	2.66
79	**12139**	0.00265	0.0155	0.00000	0.428	0.306	2.66
80	**12295**	0.00264	0.0155	0.00000	0.430	0.305	2.61
81	**12444**	0.00264	0.0155	0.00000	0.423	0.305	2.61
82	**12602**	0.00264	0.0155	0.00000	0.425	0.306	2.62
83	**12747**	0.00264	0.0155	0.00000	0.425	0.305	2.62
84	**12890**	0.00264	0.0155	0.00000	0.432	0.305	2.57
85	**13051**	0.00264	0.0155	0.00000	0.432	0.305	2.57
86	**13218**	0.00264	0.0155	0.00000	0.432	0.305	2.57
87	**13363**	0.00263	0.0155	0.00000	0.435	0.305	2.52
88	**13518**	0.00263	0.0155	0.00000	0.435	0.305	2.52
89	**13673**	0.00264	0.0155	0.00000	0.432	0.305	2.53
90	**13827**	0.00263	0.0155	0.00000	0.432	0.305	2.53
91	**13973**	0.00263	0.0154	0.00000	0.430	0.305	2.53
92	**14130**	0.00263	0.0154	0.00000	0.423	0.305	2.53
93	**14280**	0.00262	0.0154	0.00000	0.418	0.305	2.53
94	**14435**	0.00262	0.0154	0.00000	0.421	0.305	2.53
95	**14583**	0.00263	0.0154	0.00000	0.425	0.305	2.53
96	**14735**	0.00263	0.0154	0.00000	0.423	0.305	2.53
97	**14894**	0.00263	0.0154	0.00000	0.421	0.305	2.53
98	**15048**	0.00263	0.0155	0.00000	0.421	0.306	2.53

Random Forest analysis of the cars dataset predicting the Make of the car

The HPFOREST Procedure

Fit Statistics							
Number of Trees	Number of Leaves	Average Square Error (Train)	Average Square Error (OOB)	Misclassification Rate (Train)	Misclassification Rate (OOB)	Log Loss (Train)	Log Loss (OOB)
99	**15207**	0.00264	0.0155	0.00000	0.425	0.306	2.53
100	**15357**	0.00263	0.0154	0.00000	0.428	0.306	2.53

Loss Reduction Variable Importance					
Variable	Number of Rules	Gini	OOB Gini	Margin	OOB Margin
Origin	362	0.080584	0.05305	-0.01596	-0.03570
DriveTrain	172	0.020453	0.00554	0.00174	-0.00988
Cylinders	75	0.006532	0.00035	0.00191	-0.00284
Type	104	0.005973	-0.00478	0.00192	-0.00577
EngineSize	1656	0.120283	-0.00690	0.12068	0.00105
Horsepower	1175	0.074693	-0.02347	0.08667	-0.00984
Invoice	841	0.050365	-0.02816	0.05090	-0.02709
MPG_City	1380	0.074877	-0.02938	0.09111	-0.01198
MPG_Highway	1685	0.099424	-0.03495	0.11388	-0.01880
Length	1748	0.100418	-0.03721	0.13152	-0.01279
MSRP	1219	0.075114	-0.04721	0.07763	-0.04267
Wheelbase	2465	0.129548	-0.05232	0.18336	-0.00824
Weight	2375	0.118419	-0.06897	0.17718	-0.02396