

# Supplementary Materials for a Review of R Neural Network Packages (with NNbenchmark): Accuracy and Ease of Use

Salsabila Mahdi\*    Akshaj Verma†    Christophe Dutang‡    Patrice Kiener§    John C. Nash¶

## Contents

<b>1</b>	<b>Additionnal materials for all packages</b>	<b>1</b>
1.1	Result for dataset <code>mDette</code> . . . . .	1
1.2	Result for dataset <code>mFriedman</code> . . . . .	3
1.3	Result for dataset <code>mIshigami</code> . . . . .	5
1.4	Result for dataset <code>mRef153</code> . . . . .	7
1.5	Result for dataset <code>uDmod1</code> . . . . .	9
1.6	Result for dataset <code>uDmod2</code> . . . . .	11
1.7	Result for dataset <code>uDreyfus1</code> . . . . .	13
1.8	Result for dataset <code>uDreyfus2</code> . . . . .	15
1.9	Result for dataset <code>uGauss1</code> . . . . .	17
1.10	Result for dataset <code>uGauss2</code> . . . . .	19
1.11	Result for dataset <code>uGauss3</code> . . . . .	21
1.12	Result for dataset <code>uNeuroOne</code> . . . . .	23
1.13	Score probabilities . . . . .	25
<b>2</b>	<b>Additionnal materials on the large dataset <code>bWoodN1</code></b>	<b>27</b>
2.1	Summary statistics for non top-5 packages . . . . .	27
2.2	Graphics for top-5 packages . . . . .	28

## 1 Additionnal materials for all packages

### 1.1 Result for dataset `mDette`

---

\*Universitas Syiah Kuala, Indonesia

†Manipal Institute of Technology, India

‡Université Paris-Dauphine, University PSL, France

§InModelia, France

¶Telfer School of Management, University of Ottawa, Canada

Table 1: Result for mDette

Package	Algorithm	RMSE min	RMSE median	RMSE D51	MAE median	WAE median	Time mean
<b>nlr</b>	<b>41. none</b>	0.1400	0.4500	0.3100	0.3497	2.7841	0.522
<b>rminer</b>	<b>45. none</b>	0.2335	0.3147	0.0812	0.2456	1.2905	0.248
<b>nnet</b>	<b>42. none</b>	0.2650	0.4735	0.2085	0.3557	2.0121	0.078
<b>validann</b>	<b>56. BFGS</b>	0.2730	0.4266	0.1536	0.3155	1.9320	1.712
	<b>57. CG</b>	0.3813	0.4231	0.0418	0.3165	1.8043	11.228
	<b>58. L-BFGS-B</b>	0.4455	1.5927	1.1472	1.1539	8.9132	1.828
	<b>59. Nelder-Mead</b>	3.1073	3.5453	0.4380	2.7197	17.3854	2.126
	<b>60. SANN</b>	3.3417	4.0522	0.7105	2.9633	19.6574	0.172
<b>MachineShop</b>	<b>32. none</b>	0.2570	1.2314	0.9744	0.9854	8.0327	0.076
<b>traineR</b>	<b>55. none</b>	0.4539	0.5799	0.1260	0.4649	2.6448	0.078
<b>radiant.model</b>	<b>44. none</b>	0.2621	0.5412	0.2791	0.4096	2.1475	0.112
<b>monmlp</b>	<b>34. BFGS</b>	0.3732	0.4512	0.0780	0.3380	1.8359	0.298
	<b>35. Nelder-Mead</b>	3.0247	3.4557	0.4310	2.5277	18.0917	1.100
<b>CaDENCE</b>	<b>12. optim</b>	0.3277	2.5664	2.2387	1.2936	17.3208	7.072
	<b>14. Rprop</b>	4.6664	5.7488	1.0824	3.4794	31.0108	17.178
	<b>13. psoptim</b>	3.1663	3.6338	0.4675	2.1362	22.3798	11.258
<b>h2o</b>	<b>24. first-order</b>	0.3696	0.3789	0.0093	0.2948	1.3228	6.274
<b>EnsembleBase</b>	<b>23. none</b>	0.8770	13.9426	13.0656	11.3013	47.5398	0.026
<b>caret</b>	<b>15. none</b>	0.3175	0.3514	0.0339	0.2681	1.8536	0.252
<b>brnn</b>	<b>11. Gauss-Newton</b>	0.4578	1.9537	1.4959	1.4572	11.8945	0.216
<b>qrnn</b>	<b>43. none</b>	0.3632	0.7514	0.3882	0.4482	6.6249	0.518
<b>RSNNS</b>	<b>51. Rprop</b>	0.7757	1.2553	0.4796	0.9246	7.6985	0.692
	<b>52. SCG</b>	0.4652	1.7312	1.2660	1.2784	7.8765	1.156
	<b>53. Backpropagation</b>	0.4789	0.5588	0.0799	0.4219	2.0582	0.638
	<b>47. BackpropChunk</b>	0.5892	0.7126	0.1234	0.5252	2.8993	0.702
	<b>48. BackpropMomentum</b>	0.6547	0.7744	0.1197	0.5909	3.1612	0.688
	<b>49. BackpropWeightDecay</b>	0.6328	0.7698	0.1370	0.5856	3.0364	0.654
	<b>46. BackpropBatch</b>	1.9746	2.0170	0.0424	1.5451	10.0256	6.752
	<b>50. Quickprop</b>	7.1667	7.3190	0.1523	6.0055	29.6111	7.460
<b>automl</b>	<b>8. adam</b>	0.4255	0.6160	0.1905	0.4710	3.2585	9.584
	<b>9. RMSprop</b>	0.4821	0.6996	0.2175	0.5006	3.8172	8.632
	<b>10. trainwpso</b>	2.7275	4.9634	2.2359	3.7904	24.2831	13.696
<b>deepnet</b>	<b>20. BP</b>	0.5308	0.6403	0.1095	0.5135	2.7237	0.648
<b>neuralnet</b>	<b>38. rprop+</b>	0.4859	0.5467	0.0608	0.4149	2.3410	3.836
	<b>37. rprop-</b>	0.5338	2.0473	1.5135	1.4437	12.5391	6.318
	<b>40. slr</b>	0.5494	0.5688	0.0194	0.4293	2.4012	6.914
	<b>39. sag</b>	2.1196	8.1656	6.0460	6.5262	36.2385	12.916
	<b>36. backprop</b>	8.1656	8.1656	0.0000	6.5262	36.2385	14.200
<b>keras</b>	<b>28. adamax</b>	0.6492	0.6952	0.0460	0.5462	4.1959	4.386
	<b>27. adam</b>	0.7615	1.0487	0.2872	0.7949	6.3699	2.068
	<b>29. nadam</b>	1.0271	1.2485	0.2214	0.9787	4.9790	3.422
	<b>26. adagrad</b>	1.5412	2.2114	0.6702	1.5982	12.7204	18.384
	<b>25. adadelta</b>	2.0733	2.3080	0.2347	1.5890	13.7080	29.372
	<b>31. sgd</b>	0.5726	2.3026	1.7300	1.6878	10.2998	8.816
	<b>30. rmsprop</b>	2.6780	3.2516	0.5736	2.3382	16.3052	1.836
<b>AMORE</b>	<b>2. ADAPTgdwm</b>	0.3972	0.4012	0.0040	0.3084	1.7312	0.184
	<b>1. ADAPTgd</b>	0.4391	0.4564	0.0173	0.3246	2.0005	0.128
	<b>4. BATCHgdwm</b>	1.8586	1.9806	0.1220	1.4990	11.2445	1.862
	<b>3. BATCHgd</b>	1.8688	1.8999	0.0311	1.5158	8.6487	1.870
<b>minpack.lm</b>	<b>33. none</b>	0.6081	0.6081	0.0000	0.4989	1.9776	0.242
<b>ANN2</b>	<b>6. rmsprop</b>	1.9463	2.0761	0.1298	1.5240	12.6858	0.206
	<b>5. adam</b>	1.7980	2.0396	0.2416	1.5178	11.5812	0.218
	<b>7. sgd</b>	1.2208	2.0228	0.8020	1.4953	8.6218	0.204
<b>deepdive</b>	<b>16. adam</b>	3.0971	3.0971	0.0000	2.0640	18.6373	0.738
	<b>19. rmsProp</b>	2.7205	2.7205	0.0000	1.8705	16.1780	0.758
	<b>18. momentum</b>	4.1990	4.1990	0.0000	3.1011	18.5512	7.434
	<b>17. gradientDescent</b>	4.4310	4.4310	0.0000	3.2628	20.7622	7.266
<b>snnR</b>	<b>54. none</b>	1.9864	1.9864	0.0000	1.5889	8.8501	0.140
<b>elmNNRepp</b>	<b>21. extremeML</b>	7.3193	7.6899	0.3706	5.9574	32.3344	0.004
<b>ELMR</b>	<b>22. extremeML</b>	6.3469	7.2310	0.8841	5.5344	32.0052	0.018

Note: Statistics over 10 runs.

## 1.2 Result for dataset mFriedman

Table 2: Result for mFriedman

Package	Algorithm	RMSE min	RMSE median	RMSE D51	MAE median	WAE median	Time mean
<b>nlr</b>	<b>41. none</b>	0.0045	0.0061	0.0016	0.0048	0.0196	0.762
<b>rminer</b>	<b>45. none</b>	0.0095	0.0112	0.0017	0.0088	0.0373	0.286
<b>nnet</b>	<b>42. none</b>	0.0091	0.0120	0.0029	0.0094	0.0404	0.102
<b>validann</b>	<b>56. BFGS</b>	0.0096	0.0688	0.0592	0.0500	0.1513	2.500
	<b>57. CG</b>	0.0163	0.0184	0.0021	0.0145	0.0580	25.592
	<b>58. L-BFGS-B</b>	0.0211	0.0256	0.0045	0.0208	0.0865	2.702
	<b>59. Nelder-Mead</b>	0.0991	0.1082	0.0091	0.0820	0.3701	6.334
	<b>60. SANN</b>	0.1414	0.1485	0.0071	0.1149	0.5629	0.204
<b>MachineShop</b>	<b>32. none</b>	0.0085	0.0116	0.0031	0.0092	0.0360	0.106
<b>traineR</b>	<b>55. none</b>	0.0112	0.0263	0.0151	0.0209	0.0954	0.094
<b>radiant.model</b>	<b>44. none</b>	0.0084	0.0150	0.0066	0.0106	0.0664	0.122
<b>monmlp</b>	<b>34. BFGS</b>	0.0132	0.0139	0.0007	0.0110	0.0465	0.308
	<b>35. Nelder-Mead</b>	0.1155	0.1219	0.0064	0.0960	0.3777	1.084
<b>CaDENCE</b>	<b>12. optim</b>	0.0160	0.0863	0.0703	0.0442	0.3640	9.226
	<b>14. Rprop</b>	0.0850	0.1295	0.0445	0.0858	0.5842	22.516
	<b>13. psoptim</b>	0.0950	0.1148	0.0198	0.0739	0.4058	12.056
<b>h2o</b>	<b>24. first-order</b>	0.0225	0.0261	0.0036	0.0204	0.0902	6.046
<b>EnsembleBase</b>	<b>23. none</b>	0.0245	0.0262	0.0017	0.0181	0.1321	0.092
<b>caret</b>	<b>15. none</b>	0.0123	0.0197	0.0074	0.0162	0.0727	0.288
<b>brnn</b>	<b>11. Gauss-Newton</b>	0.0046	0.0052	0.0006	0.0043	0.0154	0.238
<b>qrnn</b>	<b>43. none</b>	0.0105	0.0296	0.0191	0.0190	0.1330	0.578
<b>RSNNS</b>	<b>51. Rprop</b>	0.0307	0.0452	0.0145	0.0374	0.1660	0.706
	<b>52. SCG</b>	0.0202	0.0218	0.0016	0.0170	0.0747	1.140
	<b>53. Backpropagation</b>	0.0420	0.0900	0.0480	0.0761	0.2168	0.694
	<b>47. BackpropChunk</b>	0.0541	0.0657	0.0116	0.0532	0.2284	0.732
	<b>48. BackpropMomentum</b>	0.0558	0.0789	0.0231	0.0582	0.2590	0.706
	<b>49. BackpropWeightDecay</b>	0.0429	0.0595	0.0166	0.0488	0.1832	0.726
	<b>46. BackpropBatch</b>	0.0434	0.0851	0.0417	0.0754	0.2084	6.886
	<b>50. Quickprop</b>	0.1664	0.1722	0.0058	0.1384	0.5541	7.502
<b>automl</b>	<b>8. adam</b>	0.0277	0.0323	0.0046	0.0250	0.1346	9.568
	<b>9. RMSprop</b>	0.0397	0.0504	0.0107	0.0399	0.2019	8.550
	<b>10. trainwpso</b>	0.1029	0.1228	0.0199	0.0976	0.3922	14.836
<b>deepnet</b>	<b>20. BP</b>	0.0396	0.0967	0.0571	0.0838	0.2139	0.664
<b>neuralnet</b>	<b>38. rprop+</b>	0.0102	0.0106	0.0004	0.0083	0.0356	5.862
	<b>37. rprop-</b>	0.0095	0.0110	0.0015	0.0085	0.0412	5.058
	<b>40. slr</b>	0.0690	0.2348	0.1658	0.1880	0.6346	12.928
	<b>39. sag</b>	0.0806	0.2348	0.1542	0.1880	0.6346	13.202
	<b>36. backprop</b>	0.2348	0.2348	0.0000	0.1880	0.6346	14.676
<b>keras</b>	<b>28. adamax</b>	0.0326	0.0395	0.0069	0.0319	0.1140	4.326
	<b>27. adam</b>	0.0636	0.0774	0.0138	0.0612	0.2686	2.160
	<b>29. nadam</b>	0.0732	0.0992	0.0260	0.0817	0.3144	2.482
	<b>26. adagrad</b>	0.0296	0.0842	0.0546	0.0747	0.2012	14.836
	<b>25. adadelta</b>	0.0257	0.0267	0.0010	0.0211	0.0948	29.424
	<b>31. sgd</b>	0.0365	0.0527	0.0162	0.0403	0.1922	4.136
	<b>30. rmsprop</b>	0.1010	0.1147	0.0137	0.0860	0.3822	2.240
<b>AMORE</b>	<b>2. ADAPTgdwm</b>	0.0439	0.0450	0.0011	0.0321	0.1788	0.178
	<b>1. ADAPTgd</b>	0.0264	0.0296	0.0032	0.0235	0.1101	0.128
	<b>4. BATCHgdwm</b>	0.0173	0.0176	0.0003	0.0138	0.0586	1.882
	<b>3. BATCHgd</b>	0.0177	0.0816	0.0639	0.0748	0.1692	1.876
<b>minpack.lm</b>	<b>33. none</b>	0.1269	0.1269	0.0000	0.1009	0.3714	0.380
<b>ANN2</b>	<b>6. rmsprop</b>	0.0250	0.0314	0.0064	0.0251	0.0945	0.226
	<b>5. adam</b>	0.0183	0.0201	0.0018	0.0166	0.0579	0.234
	<b>7. sgd</b>	0.0178	0.0185	0.0007	0.0147	0.0603	0.222
<b>deepdive</b>	<b>16. adam</b>	0.0875	0.0875	0.0000	0.0764	0.2699	0.768
	<b>19. rmsProp</b>	0.1287	0.1287	0.0000	0.0990	0.4133	0.772
	<b>18. momentum</b>	0.1363	0.1363	0.0000	0.1061	0.4860	7.848
	<b>17. gradientDescent</b>	0.1474	0.1474	0.0000	0.1154	0.4581	7.642
<b>snnR</b>	<b>54. none</b>	0.0457	0.0839	0.0382	0.0747	0.2113	0.102
<b>elmNNRepp</b>	<b>21. extremeML</b>	0.1516	0.1734	0.0218	0.1379	0.5055	0.000
<b>ELMR</b>	<b>22. extremeML</b>	0.1677	0.1924	0.0247	0.1538	0.5716	0.008

Note: Statistics over 10 runs.

### 1.3 Result for dataset mIshigami

Table 3: Result for mIshigami

Package	Algorithm	RMSE min	RMSE median	RMSE D51	MAE median	WAE median	Time mean
<b>nlr</b>	<b>41. none</b>	0.6602	2.2311	1.5709	1.8053	5.7864	1.470
<b>rminer</b>	<b>45. none</b>	0.6490	0.6668	0.0178	0.5016	3.0019	0.446
<b>nnet</b>	<b>42. none</b>	0.5462	0.6959	0.1497	0.5147	3.0034	0.152
<b>validann</b>	<b>56. BFGS</b>	0.6342	0.7284	0.0942	0.5216	3.3533	5.086
	<b>57. CG</b>	0.6427	0.7212	0.0785	0.5352	3.3323	58.524
	<b>58. L-BFGS-B</b>	0.8502	1.1103	0.2601	0.8812	3.5016	5.418
	<b>59. Nelder-Mead</b>	2.6029	2.6812	0.0783	2.2886	7.2908	16.114
	<b>60. SANN</b>	2.9199	2.9986	0.0787	2.4922	10.0706	0.270
<b>MachineShop</b>	<b>32. none</b>	0.6685	2.1956	1.5271	1.7857	5.3089	0.152
<b>traineR</b>	<b>55. none</b>	0.6846	0.7400	0.0554	0.5453	3.2851	0.152
<b>radiant.model</b>	<b>44. none</b>	0.4934	0.7868	0.2934	0.5896	3.1250	0.172
<b>monmlp</b>	<b>34. BFGS</b>	0.8185	0.9739	0.1554	0.7577	3.6164	0.460
	<b>35. Nelder-Mead</b>	2.7368	2.8463	0.1095	2.3257	8.7509	1.600
<b>CaDENCE</b>	<b>12. optim</b>	1.0465	1.6993	0.6528	0.8815	5.3208	14.912
	<b>14. Rprop</b>	1.3422	2.3133	0.9711	1.3927	8.8022	36.926
	<b>13. psoptim</b>	2.6775	2.7432	0.0657	2.3281	8.8488	14.936
<b>h2o</b>	<b>24. first-order</b>	0.8347	0.8467	0.0120	0.6295	3.6234	6.462
<b>EnsembleBase</b>	<b>23. none</b>	0.6342	0.8141	0.1799	0.5735	3.9523	0.132
<b>caret</b>	<b>15. none</b>	1.0310	1.6339	0.6029	1.3615	4.7983	0.418
<b>brnn</b>	<b>11. Gauss-Newton</b>	0.6588	0.6635	0.0047	0.5100	2.9395	0.204
<b>qrnn</b>	<b>43. none</b>	0.7656	0.7907	0.0251	0.4951	4.0838	1.122
<b>RSNNS</b>	<b>51. Rprop</b>	1.3146	2.3451	1.0305	1.8953	6.5010	0.840
	<b>52. SCG</b>	0.6980	0.7363	0.0383	0.5439	3.0529	1.456
	<b>53. Backpropagation</b>	2.7659	2.8040	0.0381	2.1912	11.0805	0.814
	<b>47. BackpropChunk</b>	1.3784	2.6226	1.2442	2.0664	8.9928	0.816
	<b>48. BackpropMomentum</b>	2.6138	2.6595	0.0457	2.1268	10.0368	0.822
	<b>49. BackpropWeightDecay</b>	1.2711	2.0728	0.8017	1.5275	7.9148	0.854
	<b>46. BackpropBatch</b>	2.6668	2.6742	0.0074	2.3004	7.1688	8.542
	<b>50. Quickprop</b>	3.4245	3.5389	0.1144	2.8752	13.1137	9.656
<b>automl</b>	<b>8. adam</b>	0.7511	0.7995	0.0484	0.6120	2.9212	9.932
	<b>9. RMSprop</b>	1.8225	2.5662	0.7437	2.1749	6.0520	8.882
	<b>10. trainwpso</b>	1.8381	2.4317	0.5936	1.9867	7.8872	25.376
<b>deepnet</b>	<b>20. BP</b>	1.0536	1.4687	0.4151	1.0190	6.8677	0.770
<b>neuralnet</b>	<b>38. rprop+</b>	0.5788	0.6650	0.0862	0.5052	2.7746	4.596
	<b>37. rprop-</b>	0.6728	0.7126	0.0398	0.5316	2.8674	1.954
	<b>40. slr</b>	0.6816	3.6898	3.0082	2.9776	13.1137	24.586
	<b>39. sag</b>	3.6898	3.6898	0.0000	2.9776	13.1137	25.218
	<b>36. backprop</b>	3.6898	3.6898	0.0000	2.9776	13.1137	23.958
<b>keras</b>	<b>28. adamax</b>	0.8307	0.8615	0.0308	0.6388	3.6379	5.302
	<b>27. adam</b>	0.9777	1.0728	0.0951	0.7886	4.0357	2.796
	<b>29. nadam</b>	1.0800	2.7592	1.6792	2.3587	8.0273	3.264
	<b>26. adagrad</b>	0.8522	2.5746	1.7224	2.1958	6.9534	31.856
	<b>25. adadelta</b>	2.4074	2.6007	0.1933	2.2281	6.9184	31.676
	<b>31. sgd</b>	2.7076	2.7302	0.0226	2.3252	7.5362	2.788
	<b>30. rmsprop</b>	2.8335	3.0118	0.1783	2.4550	9.4367	1.924
<b>AMORE</b>	<b>2. ADAPTgdwm</b>	0.8636	0.9950	0.1314	0.7280	3.8394	0.330
	<b>1. ADAPTgd</b>	0.7690	0.8135	0.0445	0.6083	2.9968	0.222
	<b>4. BATCHgdwm</b>	2.4805	2.5259	0.0454	2.1518	6.4536	2.678
	<b>3. BATCHgd</b>	2.5215	2.5544	0.0329	2.1768	6.3018	2.624
<b>minpack.lm</b>	<b>33. none</b>	2.5379	2.5379	0.0000	2.0524	7.6035	0.940
<b>ANN2</b>	<b>6. rmsprop</b>	0.7045	0.8590	0.1545	0.6409	2.9940	1.048
	<b>5. adam</b>	0.7560	0.8062	0.0502	0.6130	3.5492	1.058
	<b>7. sgd</b>	0.7787	0.9097	0.1310	0.6798	3.8085	1.048
<b>deepdive</b>	<b>16. adam</b>	2.5913	2.5913	0.0000	2.0819	10.0604	0.902
	<b>19. rmsProp</b>	2.6728	2.6728	0.0000	2.3060	7.1452	0.892
	<b>18. momentum</b>	2.5791	2.5791	0.0000	2.0107	8.7569	9.220
	<b>17. gradientDescent</b>	3.0218	3.0218	0.0000	2.4940	10.2360	9.062
<b>snnR</b>	<b>54. none</b>	0.7757	0.8621	0.0864	0.6030	3.4730	0.430
<b>elmNNRepp</b>	<b>21. extremeML</b>	3.0949	3.2590	0.1641	2.6511	11.3823	0.000
<b>ELMR</b>	<b>22. extremeML</b>	3.2348	3.2840	0.0492	2.6674	12.0160	0.008

Note: Statistics over 10 runs.

## 1.4 Result for dataset mRef153

Table 4: Result for mRef153

Package	Algorithm	RMSE min	RMSE median	RMSE D51	MAE median	WAE median	Time mean
<b>nlr</b>	<b>41. none</b>	0.6602	2.2311	1.5709	1.8053	5.7864	1.470
<b>rminer</b>	<b>45. none</b>	0.6490	0.6668	0.0178	0.5016	3.0019	0.446
<b>nnet</b>	<b>42. none</b>	0.5462	0.6959	0.1497	0.5147	3.0034	0.152
<b>validann</b>	<b>56. BFGS</b>	0.6342	0.7284	0.0942	0.5216	3.3533	5.086
	<b>57. CG</b>	0.6427	0.7212	0.0785	0.5352	3.3323	58.524
	<b>58. L-BFGS-B</b>	0.8502	1.1103	0.2601	0.8812	3.5016	5.418
	<b>59. Nelder-Mead</b>	2.6029	2.6812	0.0783	2.2886	7.2908	16.114
	<b>60. SANN</b>	2.9199	2.9986	0.0787	2.4922	10.0706	0.270
<b>MachineShop</b>	<b>32. none</b>	0.6685	2.1956	1.5271	1.7857	5.3089	0.152
<b>traineR</b>	<b>55. none</b>	0.6846	0.7400	0.0554	0.5453	3.2851	0.152
<b>radiant.model</b>	<b>44. none</b>	0.4934	0.7868	0.2934	0.5896	3.1250	0.172
<b>monmlp</b>	<b>34. BFGS</b>	0.8185	0.9739	0.1554	0.7577	3.6164	0.460
	<b>35. Nelder-Mead</b>	2.7368	2.8463	0.1095	2.3257	8.7509	1.600
<b>CaDENCE</b>	<b>12. optim</b>	1.0465	1.6993	0.6528	0.8815	5.3208	14.912
	<b>14. Rprop</b>	1.3422	2.3133	0.9711	1.3927	8.8022	36.926
	<b>13. psoptim</b>	2.6775	2.7432	0.0657	2.3281	8.8488	14.936
<b>h2o</b>	<b>24. first-order</b>	0.8347	0.8467	0.0120	0.6295	3.6234	6.462
<b>EnsembleBase</b>	<b>23. none</b>	0.6342	0.8141	0.1799	0.5735	3.9523	0.132
<b>caret</b>	<b>15. none</b>	1.0310	1.6339	0.6029	1.3615	4.7983	0.418
<b>brnn</b>	<b>11. Gauss-Newton</b>	0.6588	0.6635	0.0047	0.5100	2.9395	0.204
<b>qrnn</b>	<b>43. none</b>	0.7656	0.7907	0.0251	0.4951	4.0838	1.122
<b>RSNNS</b>	<b>51. Rprop</b>	1.3146	2.3451	1.0305	1.8953	6.5010	0.840
	<b>52. SCG</b>	0.6980	0.7363	0.0383	0.5439	3.0529	1.456
	<b>53. Backpropagation</b>	2.7659	2.8040	0.0381	2.1912	11.0805	0.814
	<b>47. BackpropChunk</b>	1.3784	2.6226	1.2442	2.0664	8.9928	0.816
	<b>48. BackpropMomentum</b>	2.6138	2.6595	0.0457	2.1268	10.0368	0.822
	<b>49. BackpropWeightDecay</b>	1.2711	2.0728	0.8017	1.5275	7.9148	0.854
	<b>46. BackpropBatch</b>	2.6668	2.6742	0.0074	2.3004	7.1688	8.542
	<b>50. Quickprop</b>	3.4245	3.5389	0.1144	2.8752	13.1137	9.656
<b>automl</b>	<b>8. adam</b>	0.7511	0.7995	0.0484	0.6120	2.9212	9.932
	<b>9. RMSprop</b>	1.8225	2.5662	0.7437	2.1749	6.0520	8.882
	<b>10. trainwpso</b>	1.8381	2.4317	0.5936	1.9867	7.8872	25.376
<b>deepnet</b>	<b>20. BP</b>	1.0536	1.4687	0.4151	1.0190	6.8677	0.770
<b>neuralnet</b>	<b>38. rprop+</b>	0.5788	0.6650	0.0862	0.5052	2.7746	4.596
	<b>37. rprop-</b>	0.6728	0.7126	0.0398	0.5316	2.8674	1.954
	<b>40. slr</b>	0.6816	3.6898	3.0082	2.9776	13.1137	24.586
	<b>39. sag</b>	3.6898	3.6898	0.0000	2.9776	13.1137	25.218
	<b>36. backprop</b>	3.6898	3.6898	0.0000	2.9776	13.1137	23.958
<b>keras</b>	<b>28. adamax</b>	0.8307	0.8615	0.0308	0.6388	3.6379	5.302
	<b>27. adam</b>	0.9777	1.0728	0.0951	0.7886	4.0357	2.796
	<b>29. nadam</b>	1.0800	2.7592	1.6792	2.3587	8.0273	3.264
	<b>26. adagrad</b>	0.8522	2.5746	1.7224	2.1958	6.9534	31.856
	<b>25. adadelta</b>	2.4074	2.6007	0.1933	2.2281	6.9184	31.676
	<b>31. sgd</b>	2.7076	2.7302	0.0226	2.3252	7.5362	2.788
	<b>30. rmsprop</b>	2.8335	3.0118	0.1783	2.4550	9.4367	1.924
<b>AMORE</b>	<b>2. ADAPTgdwm</b>	0.8636	0.9950	0.1314	0.7280	3.8394	0.330
	<b>1. ADAPTgd</b>	0.7690	0.8135	0.0445	0.6083	2.9968	0.222
	<b>4. BATCHgdwm</b>	2.4805	2.5259	0.0454	2.1518	6.4536	2.678
	<b>3. BATCHgd</b>	2.5215	2.5544	0.0329	2.1768	6.3018	2.624
<b>minpack.lm</b>	<b>33. none</b>	2.5379	2.5379	0.0000	2.0524	7.6035	0.940
<b>ANN2</b>	<b>6. rmsprop</b>	0.7045	0.8590	0.1545	0.6409	2.9940	1.048
	<b>5. adam</b>	0.7560	0.8062	0.0502	0.6130	3.5492	1.058
	<b>7. sgd</b>	0.7787	0.9097	0.1310	0.6798	3.8085	1.048
<b>deepdive</b>	<b>16. adam</b>	2.5913	2.5913	0.0000	2.0819	10.0604	0.902
	<b>19. rmsProp</b>	2.6728	2.6728	0.0000	2.3060	7.1452	0.892
	<b>18. momentum</b>	2.5791	2.5791	0.0000	2.0107	8.7569	9.220
	<b>17. gradientDescent</b>	3.0218	3.0218	0.0000	2.4940	10.2360	9.062
<b>snnR</b>	<b>54. none</b>	0.7757	0.8621	0.0864	0.6030	3.4730	0.430
<b>elmNNRepp</b>	<b>21. extremeML</b>	3.0949	3.2590	0.1641	2.6511	11.3823	0.000
<b>ELMR</b>	<b>22. extremeML</b>	3.2348	3.2840	0.0492	2.6674	12.0160	0.008

Note: Statistics over 10 runs.



## 1.5 Result for dataset uDmod1

Table 5: Result for uDmod1

Package	Algorithm	RMSE min	RMSE median	RMSE D51	MAE median	WAE median	Time mean
<b>nlr</b>	<b>41. none</b>	0.0433	0.0433	0.0000	0.0349	0.1063	0.088
<b>rminer</b>	<b>45. none</b>	0.0449	0.0495	0.0046	0.0418	0.1258	0.030
<b>nnet</b>	<b>42. none</b>	0.0437	0.0865	0.0428	0.0636	0.3435	0.008
<b>validann</b>	<b>56. BFGS</b>	0.0435	0.0725	0.0290	0.0540	0.1810	0.790
	<b>57. CG</b>	0.0506	0.0679	0.0173	0.0544	0.1577	29.066
	<b>58. L-BFGS-B</b>	0.0489	0.1090	0.0601	0.0759	0.4093	0.880
	<b>59. Nelder-Mead</b>	0.1034	0.1810	0.0776	0.1538	0.4017	28.208
	<b>60. SANN</b>	0.2296	0.3046	0.0750	0.2441	0.6614	0.128
<b>MachineShop</b>	<b>32. none</b>	0.0442	0.0456	0.0014	0.0365	0.1181	0.012
<b>traineR</b>	<b>55. none</b>	0.0410	0.0470	0.0060	0.0393	0.1293	0.004
<b>radiant.model</b>	<b>44. none</b>	0.0800	0.1088	0.0288	0.0817	0.3346	0.026
<b>monmlp</b>	<b>34. BFGS</b>	0.0919	0.0983	0.0064	0.0750	0.3693	0.208
	<b>35. Nelder-Mead</b>	0.1381	0.2639	0.1258	0.2153	0.6177	0.428
<b>CaDENCE</b>	<b>12. optim</b>	0.0564	0.2112	0.1548	0.1061	0.6888	2.442
	<b>14. Rprop</b>	0.2005	0.4116	0.2111	0.3162	0.8665	6.654
	<b>13. psoptim</b>	0.3096	0.3190	0.0094	0.2672	0.7427	5.378
<b>h2o</b>	<b>24. first-order</b>	0.0480	0.0494	0.0014	0.0402	0.1185	3.346
<b>EnsembleBase</b>	<b>23. none</b>	0.0733	0.1033	0.0300	0.0759	0.4193	0.004
<b>caret</b>	<b>15. none</b>	0.0535	0.0948	0.0413	0.0602	0.3176	0.030
<b>brnn</b>	<b>11. Gauss-Newton</b>	0.0451	0.5884	0.5433	0.5069	1.0104	0.010
<b>qrnn</b>	<b>43. none</b>	0.1162	0.1349	0.0187	0.0830	0.6014	0.230
<b>RSNNS</b>	<b>51. Rprop</b>	0.1232	0.1401	0.0169	0.1048	0.4453	0.090
	<b>52. SCG</b>	0.0970	0.1118	0.0148	0.0916	0.4280	0.140
	<b>53. Backpropagation</b>	0.1215	0.2226	0.1011	0.1736	0.5618	0.094
	<b>47. BackpropChunk</b>	0.1298	0.1448	0.0150	0.1073	0.5245	0.140
	<b>48. BackpropMomentum</b>	0.1445	0.1647	0.0202	0.1252	0.5800	0.088
	<b>49. BackpropWeightDecay</b>	0.1314	0.1656	0.0342	0.1218	0.5395	0.090
	<b>46. BackpropBatch</b>	0.2568	0.3344	0.0776	0.2870	0.7691	0.874
	<b>50. Quickprop</b>	0.5775	0.5884	0.0109	0.5068	1.0104	0.938
<b>automl</b>	<b>8. adam</b>	0.0596	0.1157	0.0561	0.0741	0.5060	1.262
	<b>9. RMSprop</b>	0.1052	0.1595	0.0543	0.1323	0.3299	1.128
	<b>10. trainwpso</b>	0.2424	0.2517	0.0093	0.1929	0.6461	6.964
<b>deepnet</b>	<b>20. BP</b>	0.0582	0.1173	0.0591	0.0845	0.3896	0.094
<b>neuralnet</b>	<b>38. rprop+</b>	0.1086	0.1639	0.0553	0.1319	0.5153	0.042
	<b>37. rprop-</b>	0.1634	0.1750	0.0116	0.1370	0.5212	0.030
	<b>40. slr</b>	0.0839	0.1213	0.0374	0.0922	0.3196	0.100
	<b>39. sag</b>	0.0583	0.1315	0.0732	0.1061	0.3669	1.430
	<b>36. backprop</b>	0.1521	0.1699	0.0178	0.1280	0.5924	0.490
<b>keras</b>	<b>28. adamax</b>	0.0883	0.2240	0.1357	0.1782	0.5854	4.566
	<b>27. adam</b>	0.1376	0.1811	0.0435	0.1461	0.4721	2.576
	<b>29. nadam</b>	0.1786	0.2607	0.0821	0.2055	0.6971	2.242
	<b>26. adagrad</b>	0.2252	0.3529	0.1277	0.3037	0.8099	8.322
	<b>25. adadelta</b>	0.2314	0.2333	0.0019	0.1843	0.5698	23.966
	<b>31. sgd</b>	0.2044	0.3548	0.1504	0.2992	0.8224	2.644
	<b>30. rmsprop</b>	0.2375	0.3800	0.1425	0.2964	0.8503	1.326
<b>AMORE</b>	<b>2. ADAPTgdwm</b>	0.2197	0.2765	0.0568	0.2204	0.6575	0.054
	<b>1. ADAPTgd</b>	0.3082	0.3271	0.0189	0.2829	0.7263	0.036
	<b>4. BATCHgdwm</b>	0.3265	0.3274	0.0009	0.2853	0.7289	1.804
	<b>3. BATCHgd</b>	0.2023	0.2922	0.0899	0.2393	0.6802	1.780
<b>minpack.lm</b>	<b>33. none</b>	0.0445	0.0445	0.0000	0.0362	0.1153	0.038
<b>ANN2</b>	<b>6. rmsprop</b>	0.2345	0.2495	0.0150	0.1926	0.6040	0.016
	<b>5. adam</b>	0.2198	0.2274	0.0076	0.1806	0.5242	0.012
	<b>7. sgd</b>	0.2581	0.3342	0.0761	0.2899	0.6824	0.014
<b>deepdive</b>	<b>16. adam</b>	0.1178	0.1178	0.0000	0.0797	0.4868	0.568
	<b>19. rmsProp</b>	0.1728	0.1728	0.0000	0.1257	0.4478	0.584
	<b>18. momentum</b>	0.3320	0.3320	0.0000	0.2891	0.7441	5.524
	<b>17. gradientDescent</b>	0.3353	0.3353	0.0000	0.2912	0.7067	5.340
<b>snnR</b>	<b>54. none</b>	0.2927	0.2927	0.0000	0.2512	0.6561	0.040
<b>elmNNRepp</b>	<b>21. extremeML</b>	0.3320	0.3623	0.0303	0.3038	0.8727	0.000
<b>ELMR</b>	<b>22. extremeML</b>	0.3003	0.3082	0.0079	0.2529	0.7867	0.000

Note: Statistics over 10 runs.

## 1.6 Result for dataset uDmod2

Table 6: Result for uDmod2

Package	Algorithm	RMSE min	RMSE median	RMSE D51	MAE median	WAE median	Time mean
<b>nlr</b>	<b>41. none</b>	0.0427	0.0427	0.0000	0.0333	0.1058	0.036
<b>rminer</b>	<b>45. none</b>	0.0405	0.0579	0.0174	0.0479	0.1065	0.016
<b>nnet</b>	<b>42. none</b>	0.0602	0.0615	0.0013	0.0489	0.1408	0.008
<b>validann</b>	<b>56. BFGS</b>	0.0405	0.0437	0.0032	0.0342	0.1131	0.676
	<b>57. CG</b>	0.0536	0.0610	0.0074	0.0476	0.1415	31.616
	<b>58. L-BFGS-B</b>	0.0691	0.0778	0.0087	0.0617	0.2019	0.784
	<b>59. Nelder-Mead</b>	0.0633	0.1987	0.1354	0.1673	0.4305	30.686
	<b>60. SANN</b>	0.2274	0.2546	0.0272	0.1990	0.5099	0.148
<b>MachineShop</b>	<b>32. none</b>	0.0406	0.0494	0.0088	0.0374	0.1288	0.012
<b>traineR</b>	<b>55. none</b>	0.0505	0.0649	0.0144	0.0529	0.1392	0.016
<b>radiant.model</b>	<b>44. none</b>	0.0647	0.0771	0.0124	0.0602	0.2202	0.022
<b>monmlp</b>	<b>34. BFGS</b>	0.0522	0.0796	0.0274	0.0625	0.2280	0.210
	<b>35. Nelder-Mead</b>	0.1342	0.1780	0.0438	0.1371	0.4534	0.370
<b>CaDENCE</b>	<b>12. optim</b>	0.0688	0.0805	0.0117	0.0582	0.2385	2.310
	<b>14. Rprop</b>	0.1820	0.2615	0.0795	0.2061	0.6887	5.784
	<b>13. psoptim</b>	0.2114	0.3238	0.1124	0.2711	0.6808	5.052
<b>h2o</b>	<b>24. first-order</b>	0.0474	0.0482	0.0008	0.0394	0.1126	3.342
<b>EnsembleBase</b>	<b>23. none</b>	0.0618	0.0638	0.0020	0.0504	0.1716	0.006
<b>caret</b>	<b>15. none</b>	0.0512	0.0558	0.0046	0.0454	0.1561	0.024
<b>brnn</b>	<b>11. Gauss-Newton</b>	0.0435	0.0673	0.0238	0.0522	0.1838	0.020
<b>qrnn</b>	<b>43. none</b>	0.0511	0.0821	0.0310	0.0598	0.2411	0.214
<b>RSNNS</b>	<b>51. Rprop</b>	0.0447	0.0959	0.0512	0.0717	0.2622	0.090
	<b>52. SCG</b>	0.0555	0.0788	0.0233	0.0618	0.2070	0.132
	<b>53. Backpropagation</b>	0.0788	0.1292	0.0504	0.0999	0.3342	0.086
	<b>47. BackpropChunk</b>	0.0829	0.0892	0.0063	0.0732	0.2035	0.090
	<b>48. BackpropMomentum</b>	0.0752	0.0964	0.0212	0.0786	0.2134	0.092
	<b>49. BackpropWeightDecay</b>	0.0799	0.0888	0.0089	0.0704	0.2063	0.094
	<b>46. BackpropBatch</b>	0.2601	0.2736	0.0135	0.2371	0.6099	0.870
	<b>50. Quickprop</b>	0.2570	0.4804	0.2234	0.4177	1.0187	0.906
<b>automl</b>	<b>8. adam</b>	0.0511	0.0867	0.0356	0.0707	0.1976	1.260
	<b>9. RMSprop</b>	0.1245	0.2296	0.1051	0.1669	0.5276	1.104
	<b>10. trainwpso</b>	0.2032	0.2573	0.0541	0.2232	0.5240	10.432
<b>deepnet</b>	<b>20. BP</b>	0.0563	0.0608	0.0045	0.0490	0.1446	0.092
<b>neuralnet</b>	<b>38. rprop+</b>	0.1077	0.1207	0.0130	0.0926	0.2648	0.036
	<b>37. rprop-</b>	0.0955	0.1186	0.0231	0.0920	0.2812	0.062
	<b>40. slr</b>	0.0840	0.1039	0.0199	0.0866	0.2554	0.092
	<b>39. sag</b>	0.0811	0.1160	0.0349	0.0943	0.2960	0.950
	<b>36. backprop</b>	0.1091	0.1355	0.0264	0.1084	0.3411	0.372
<b>keras</b>	<b>28. adamax</b>	0.1082	0.1728	0.0646	0.1393	0.3885	3.696
	<b>27. adam</b>	0.0963	0.1767	0.0804	0.1397	0.4913	2.338
	<b>29. nadam</b>	0.1201	0.1884	0.0683	0.1486	0.5214	2.460
	<b>26. adagrad</b>	0.1597	0.1792	0.0195	0.1379	0.4143	13.616
	<b>25. adadelta</b>	0.1746	0.1792	0.0046	0.1379	0.4116	26.680
	<b>31. sgd</b>	0.2431	0.3056	0.0625	0.2606	0.6923	1.868
	<b>30. rmsprop</b>	0.1629	0.2166	0.0537	0.1697	0.5175	1.744
<b>AMORE</b>	<b>2. ADAPTgdwm</b>	0.1145	0.1924	0.0779	0.1573	0.4195	0.034
	<b>1. ADAPTgd</b>	0.2579	0.2632	0.0053	0.2333	0.5307	0.022
	<b>4. BATCHgdwm</b>	0.1585	0.2621	0.1036	0.2341	0.4898	1.650
	<b>3. BATCHgd</b>	0.2228	0.2644	0.0416	0.2347	0.4989	1.650
<b>minpack.lm</b>	<b>33. none</b>	0.0427	0.0427	0.0000	0.0333	0.1058	0.024
<b>ANN2</b>	<b>6. rmsprop</b>	0.1831	0.2585	0.0754	0.2227	0.5514	0.012
	<b>5. adam</b>	0.1702	0.2126	0.0424	0.1747	0.4630	0.014
	<b>7. sgd</b>	0.2518	0.2732	0.0214	0.2401	0.5272	0.012
<b>deepdive</b>	<b>16. adam</b>	0.3189	0.3189	0.0000	0.2113	0.7209	0.572
	<b>19. rmsProp</b>	0.2252	0.2252	0.0000	0.1580	0.5513	0.558
	<b>18. momentum</b>	0.2656	0.2656	0.0000	0.2355	0.5269	5.460
	<b>17. gradientDescent</b>	0.2699	0.2699	0.0000	0.2369	0.5509	5.326
<b>snnR</b>	<b>54. none</b>	0.2585	0.2984	0.0399	0.2556	0.6651	0.020
<b>elmNNRepp</b>	<b>21. extremeML</b>	0.2589	0.2648	0.0059	0.2308	0.5419	0.000
<b>ELMR</b>	<b>22. extremeML</b>	0.2613	0.2735	0.0122	0.2358	0.5366	0.000

Note: Statistics over 10 runs.

## 1.7 Result for dataset uDreyfus1

Table 7: Result for uDreyfus1

Package	Algorithm	RMSE min	RMSE median	RMSE D51	MAE median	WAE median	Time mean
<b>nlr</b>	<b>41. none</b>	0.0000	0.0000	0.0000	0.0000	0.0001	0.014
<b>rminer</b>	<b>45. none</b>	0.0020	0.0023	0.0003	0.0018	0.0057	0.012
<b>nnet</b>	<b>42. none</b>	0.0026	0.0716	0.0690	0.0449	0.2254	0.004
<b>validann</b>	<b>56. BFGS</b>	0.0022	0.0023	0.0001	0.0019	0.0070	0.350
	<b>57. CG</b>	0.0035	0.0076	0.0041	0.0061	0.0205	25.106
	<b>58. L-BFGS-B</b>	0.0038	0.0084	0.0046	0.0066	0.0207	0.512
	<b>59. Nelder-Mead</b>	0.0833	0.1951	0.1118	0.1633	0.3615	17.878
	<b>60. SANN</b>	0.2692	0.3271	0.0579	0.2707	0.6914	0.142
<b>MachineShop</b>	<b>32. none</b>	0.0023	0.0034	0.0011	0.0028	0.0102	0.010
<b>traineR</b>	<b>55. none</b>	0.0019	0.0022	0.0003	0.0019	0.0076	0.000
<b>radiant.model</b>	<b>44. none</b>	0.0121	0.0682	0.0561	0.0555	0.1546	0.022
<b>monmlp</b>	<b>34. BFGS</b>	0.0323	0.0541	0.0218	0.0434	0.1524	0.190
	<b>35. Nelder-Mead</b>	0.1425	0.2017	0.0592	0.1653	0.4572	0.270
<b>CaDENCE</b>	<b>12. optim</b>	0.0032	0.6701	0.6669	0.3727	1.9004	1.030
	<b>14. Rprop</b>	0.3995	1.1290	0.7295	0.8219	2.2557	3.762
	<b>13. psoptim</b>	0.4218	0.5720	0.1502	0.3001	1.6829	4.512
<b>h2o</b>	<b>24. first-order</b>	0.0131	0.0146	0.0015	0.0112	0.0432	3.334
<b>EnsembleBase</b>	<b>23. none</b>	0.0922	0.1151	0.0229	0.0834	0.3335	0.002
<b>caret</b>	<b>15. none</b>	0.0262	0.0359	0.0097	0.0282	0.1115	0.018
<b>brnn</b>	<b>11. Gauss-Newton</b>	0.0026	0.0034	0.0008	0.0029	0.0115	0.000
<b>qrnn</b>	<b>43. none</b>	0.2781	0.2841	0.0060	0.1815	0.9095	0.128
<b>RSNNS</b>	<b>51. Rprop</b>	0.0617	0.0689	0.0072	0.0484	0.2211	0.080
	<b>52. SCG</b>	0.0851	0.1018	0.0167	0.0848	0.2408	0.122
	<b>53. Backpropagation</b>	0.1127	0.1190	0.0063	0.1000	0.2547	0.078
	<b>47. BackpropChunk</b>	0.0838	0.1275	0.0437	0.0822	0.3313	0.088
	<b>48. BackpropMomentum</b>	0.0719	0.0795	0.0076	0.0606	0.2070	0.080
	<b>49. BackpropWeightDecay</b>	0.0797	0.0849	0.0052	0.0657	0.2541	0.090
	<b>46. BackpropBatch</b>	0.3120	0.3387	0.0267	0.2647	0.7662	0.812
	<b>50. Quickprop</b>	0.2177	0.2408	0.0231	0.2084	0.5154	0.822
<b>automl</b>	<b>8. adam</b>	0.0087	0.0725	0.0638	0.0481	0.2070	0.906
	<b>9. RMSprop</b>	0.0479	0.0727	0.0248	0.0498	0.2335	1.114
	<b>10. trainwpso</b>	0.1052	0.1154	0.0102	0.0854	0.3281	5.358
<b>deepnet</b>	<b>20. BP</b>	0.0139	0.0704	0.0565	0.0451	0.2219	0.084
<b>neuralnet</b>	<b>38. rprop+</b>	0.2119	0.3475	0.1356	0.2662	0.7910	0.004
	<b>37. rprop-</b>	0.1014	0.2856	0.1842	0.2227	0.7157	0.008
	<b>40. slr</b>	0.2981	0.3450	0.0469	0.2730	0.7821	0.012
	<b>39. sag</b>	0.1963	0.3371	0.1408	0.2652	0.7510	0.048
	<b>36. backprop</b>	0.3201	0.3503	0.0302	0.2743	0.7831	0.040
<b>keras</b>	<b>28. adamax</b>	0.0365	0.0487	0.0122	0.0404	0.1489	5.074
	<b>27. adam</b>	0.0706	0.0897	0.0191	0.0690	0.2075	2.808
	<b>29. nadam</b>	0.0648	0.1550	0.0902	0.1179	0.3970	2.082
	<b>26. adagrad</b>	0.1630	0.3528	0.1898	0.2697	0.8045	5.918
	<b>25. adadelta</b>	0.2178	0.3498	0.1320	0.2655	0.8040	10.958
	<b>31. sgd</b>	0.3373	0.3450	0.0077	0.2698	0.7744	2.312
	<b>30. rmsprop</b>	0.3101	0.3622	0.0521	0.2798	0.9265	0.942
<b>AMORE</b>	<b>2. ADAPTgdwm</b>	0.1804	0.2112	0.0308	0.1476	0.4856	0.030
	<b>1. ADAPTgd</b>	0.3308	0.3475	0.0167	0.2718	0.7716	0.020
	<b>4. BATCHgdwm</b>	0.3346	0.3370	0.0024	0.2785	0.7142	1.376
	<b>3. BATCHgd</b>	0.3160	0.3346	0.0186	0.2740	0.7087	1.382
<b>minpack.lm</b>	<b>33. none</b>	0.0000	0.0000	0.0000	0.0000	0.0001	0.000
<b>ANN2</b>	<b>6. rmsprop</b>	0.2467	0.3428	0.0961	0.2715	0.7616	0.006
	<b>5. adam</b>	0.2762	0.3201	0.0439	0.2542	0.7372	0.006
	<b>7. sgd</b>	0.3493	0.3546	0.0053	0.2689	0.8481	0.008
<b>deepdive</b>	<b>16. adam</b>	0.0304	0.0304	0.0000	0.0265	0.0644	0.564
	<b>19. rmsProp</b>	0.1184	0.1184	0.0000	0.0878	0.3446	0.566
	<b>18. momentum</b>	0.3429	0.3429	0.0000	0.2801	0.7341	5.460
	<b>17. gradientDescent</b>	0.3429	0.3429	0.0000	0.2801	0.7346	5.162
<b>snnR</b>	<b>54. none</b>	0.3691	0.3691	0.0000	0.2756	0.8531	0.006
<b>elmNNRepp</b>	<b>21. extremeML</b>	0.3407	0.4066	0.0659	0.2973	1.0342	0.000
<b>ELMR</b>	<b>22. extremeML</b>	0.3987	0.4505	0.0518	0.3027	1.1845	0.000

Note: Statistics over 10 runs.

## 1.8 Result for dataset uDreyfus2

Table 8: Result for uDreyfus2

Package	Algorithm	RMSE min	RMSE median	RMSE D51	MAE median	WAE median	Time mean
<b>nlr</b>	<b>41. none</b>	0.0906	0.0906	0.0000	0.0723	0.2197	0.064
<b>rminer</b>	<b>45. none</b>	0.0906	0.0906	0.0000	0.0724	0.2202	0.018
<b>nnet</b>	<b>42. none</b>	0.0906	0.0906	0.0000	0.0725	0.2202	0.000
<b>validann</b>	<b>56. BFGS</b>	0.0906	0.0906	0.0000	0.0724	0.2200	0.442
	<b>57. CG</b>	0.0910	0.0913	0.0003	0.0730	0.2244	26.808
	<b>58. L-BFGS-B</b>	0.0907	0.1123	0.0216	0.0897	0.2733	0.504
	<b>59. Nelder-Mead</b>	0.1300	0.1604	0.0304	0.1224	0.4798	16.024
	<b>60. SANN</b>	0.2712	0.2972	0.0260	0.2354	0.7465	0.152
<b>MachineShop</b>	<b>32. none</b>	0.0906	0.1415	0.0509	0.1045	0.4507	0.012
<b>traineR</b>	<b>55. none</b>	0.0906	0.1123	0.0217	0.0901	0.2736	0.004
<b>radiant.model</b>	<b>44. none</b>	0.0907	0.0917	0.0010	0.0732	0.2275	0.026
<b>monmlp</b>	<b>34. BFGS</b>	0.0917	0.0951	0.0034	0.0746	0.2363	0.210
	<b>35. Nelder-Mead</b>	0.1762	0.2448	0.0686	0.1940	0.6268	0.248
<b>CaDENCE</b>	<b>12. optim</b>	0.0924	0.3856	0.2932	0.2560	1.1222	1.046
	<b>14. Rprop</b>	0.1684	0.2586	0.0902	0.1963	0.7853	4.272
	<b>13. psoptim</b>	0.3210	0.3814	0.0604	0.2825	1.0638	4.500
<b>h2o</b>	<b>24. first-order</b>	0.0926	0.0933	0.0007	0.0740	0.2242	3.356
<b>EnsembleBase</b>	<b>23. none</b>	0.1196	0.1272	0.0076	0.1022	0.3102	0.008
<b>caret</b>	<b>15. none</b>	0.0926	0.1039	0.0113	0.0811	0.2375	0.022
<b>brnn</b>	<b>11. Gauss-Newton</b>	0.0913	0.0913	0.0000	0.0730	0.2241	0.000
<b>qrnn</b>	<b>43. none</b>	0.1601	0.2693	0.1092	0.1983	0.6891	0.170
<b>RSNNS</b>	<b>51. Rprop</b>	0.1145	0.1252	0.0107	0.1009	0.3338	0.082
	<b>52. SCG</b>	0.1238	0.2542	0.1304	0.1876	0.7205	0.114
	<b>53. Backpropagation</b>	0.1298	0.1325	0.0027	0.1025	0.3352	0.080
	<b>47. BackpropChunk</b>	0.1199	0.1689	0.0490	0.1245	0.5094	0.084
	<b>48. BackpropMomentum</b>	0.1209	0.1297	0.0088	0.1042	0.3327	0.086
	<b>49. BackpropWeightDecay</b>	0.1186	0.1214	0.0028	0.0950	0.2751	0.082
	<b>46. BackpropBatch</b>	0.3063	0.3491	0.0428	0.2736	0.8922	0.820
	<b>50. Quickprop</b>	0.2122	0.2993	0.0871	0.2338	0.8131	0.826
<b>automl</b>	<b>8. adam</b>	0.0933	0.1579	0.0646	0.1212	0.4579	1.244
	<b>9. RMSprop</b>	0.1179	0.1615	0.0436	0.1223	0.4550	1.096
	<b>10. trainwpso</b>	0.1180	0.1616	0.0436	0.1233	0.4403	5.164
<b>deepnet</b>	<b>20. BP</b>	0.0928	0.1049	0.0121	0.0824	0.2577	0.080
<b>neuralnet</b>	<b>38. rprop+</b>	0.2846	0.3562	0.0716	0.2762	0.8861	0.008
	<b>37. rprop-</b>	0.1632	0.3537	0.1905	0.2756	0.9038	0.014
	<b>40. slr</b>	0.3374	0.3435	0.0061	0.2714	0.9130	0.012
	<b>39. sag</b>	0.1663	0.2521	0.0858	0.1997	0.7213	0.098
	<b>36. backprop</b>	0.3205	0.3655	0.0450	0.2776	0.9575	0.042
<b>keras</b>	<b>28. adamax</b>	0.1118	0.1154	0.0036	0.0935	0.3048	4.740
	<b>27. adam</b>	0.1105	0.1345	0.0240	0.1062	0.3156	2.354
	<b>29. nadam</b>	0.1341	0.1957	0.0616	0.1530	0.5132	2.366
	<b>26. adagrad</b>	0.1847	0.1979	0.0132	0.1420	0.6153	14.992
	<b>25. adadelta</b>	0.3605	0.3726	0.0121	0.2763	0.9935	7.248
	<b>31. sgd</b>	0.3523	0.3548	0.0025	0.2760	0.9224	2.546
	<b>30. rmsprop</b>	0.2221	0.3598	0.1377	0.2805	0.8788	1.096
<b>AMORE</b>	<b>2. ADAPTgdwm</b>	0.1675	0.2519	0.0844	0.1964	0.6689	0.030
	<b>1. ADAPTgd</b>	0.3555	0.3612	0.0057	0.2824	0.9010	0.020
	<b>4. BATCHgdwm</b>	0.2097	0.3405	0.1308	0.2704	0.8640	1.380
	<b>3. BATCHgd</b>	0.1778	0.2708	0.0930	0.2148	0.6304	1.378
<b>minpack.lm</b>	<b>33. none</b>	0.0906	0.0906	0.0000	0.0723	0.2197	0.022
<b>ANN2</b>	<b>6. rmsprop</b>	0.2338	0.2845	0.0507	0.2093	0.8061	0.014
	<b>5. adam</b>	0.3222	0.3836	0.0614	0.2918	0.8958	0.010
	<b>7. sgd</b>	0.3581	0.3717	0.0136	0.2806	0.9555	0.008
<b>deepdive</b>	<b>16. adam</b>	0.1149	0.1149	0.0000	0.0907	0.2749	0.566
	<b>19. rmsProp</b>	0.1625	0.1625	0.0000	0.1232	0.4839	0.552
	<b>18. momentum</b>	0.3570	0.3570	0.0000	0.2907	0.8468	5.390
	<b>17. gradientDescent</b>	0.3570	0.3570	0.0000	0.2905	0.8478	5.222
<b>snnR</b>	<b>54. none</b>	0.3837	0.3837	0.0000	0.2773	1.0352	0.012
<b>elmNNRepp</b>	<b>21. extremeML</b>	0.4534	0.6226	0.1692	0.5077	1.4031	0.000
<b>ELMR</b>	<b>22. extremeML</b>	0.4554	0.5844	0.1290	0.4293	1.3727	0.002

Note: Statistics over 10 runs.



## 1.9 Result for dataset uGauss1

Table 9: Result for uGauss1

Package	Algorithm	RMSE min	RMSE median	RMSE D51	MAE median	WAE median	Time mean
<b>nlr</b>	<b>41. none</b>	2.2321	2.6129	0.3808	2.0484	8.2482	0.156
<b>rminer</b>	<b>45. none</b>	2.2452	2.2628	0.0176	1.7485	7.3946	0.084
<b>nnet</b>	<b>42. none</b>	2.2380	2.2557	0.0177	1.7541	7.3381	0.032
<b>validann</b>	<b>56. BFGS</b>	2.2606	2.3192	0.0586	1.8215	7.5409	0.932
	<b>57. CG</b>	2.3620	2.3962	0.0342	1.9105	8.2206	41.996
	<b>58. L-BFGS-B</b>	2.7065	3.4311	0.7246	2.6492	9.8990	1.060
	<b>59. Nelder-Mead</b>	9.2009	11.1644	1.9635	9.1979	26.4353	42.914
	<b>60. SANN</b>	12.6829	15.2562	2.5733	12.8000	37.0901	0.204
<b>MachineShop</b>	<b>32. none</b>	2.2521	2.2681	0.0160	1.7734	7.5718	0.034
<b>traineR</b>	<b>55. none</b>	2.2431	2.3022	0.0591	1.7999	7.5122	0.038
<b>radiant.model</b>	<b>44. none</b>	2.3275	6.3010	3.9735	5.1479	17.4222	0.054
<b>monmlp</b>	<b>34. BFGS</b>	2.7246	5.7229	2.9983	4.8218	14.7964	0.226
	<b>35. Nelder-Mead</b>	12.1868	12.7777	0.5909	9.8127	35.0537	0.574
<b>CaDENCE</b>	<b>12. optim</b>	2.3392	2.4124	0.0732	1.9079	7.6898	2.906
	<b>14. Rprop</b>	17.7666	25.2780	7.5114	17.3441	56.1202	9.070
	<b>13. psoptim</b>	25.3213	29.0788	3.7575	22.9938	64.9998	6.144
<b>h2o</b>	<b>24. first-order</b>	2.2985	2.3587	0.0602	1.8349	7.6990	4.372
<b>EnsembleBase</b>	<b>23. none</b>	2.5261	2.6857	0.1596	2.0641	8.3968	0.038
<b>caret</b>	<b>15. none</b>	2.3241	2.6365	0.3124	2.0948	7.8111	0.108
<b>brnn</b>	<b>11. Gauss-Newton</b>	2.2434	2.4366	0.1932	1.8814	8.7824	0.042
<b>qrnn</b>	<b>43. none</b>	2.7155	2.7208	0.0053	2.0888	8.1916	0.158
<b>RSNNS</b>	<b>51. Rprop</b>	2.9859	10.3291	7.3432	7.2486	33.1339	0.320
	<b>52. SCG</b>	2.7001	4.7483	2.0482	3.6042	16.9268	0.558
	<b>53. Backpropagation</b>	3.0594	3.2011	0.1417	2.5444	9.0758	0.370
	<b>47. BackpropChunk</b>	2.9425	2.9820	0.0395	2.4160	10.5110	0.368
	<b>48. BackpropMomentum</b>	2.8791	2.9074	0.0283	2.2991	9.6946	0.334
	<b>49. BackpropWeightDecay</b>	2.8682	3.0805	0.2123	2.3657	8.8148	0.372
	<b>46. BackpropBatch</b>	14.5624	19.9070	5.3446	16.5809	52.1745	3.478
	<b>50. Quickprop</b>	23.6323	24.1323	0.5000	20.1405	57.6774	3.696
<b>automl</b>	<b>8. adam</b>	4.4523	4.9986	0.5463	3.8522	18.0133	4.982
	<b>9. RMSprop</b>	4.9906	5.2304	0.2398	4.1778	17.4681	4.482
	<b>10. trainwpso</b>	9.8111	13.4226	3.6115	9.5187	39.8505	8.586
<b>deepnet</b>	<b>20. BP</b>	3.4191	4.2839	0.8648	3.5260	11.8002	0.304
<b>neuralnet</b>	<b>38. rprop+</b>	2.9603	3.7997	0.8394	2.8334	14.1152	0.298
	<b>37. rprop-</b>	2.6198	3.8778	1.2580	2.9818	13.8198	0.308
	<b>40. slr</b>	2.8881	4.3263	1.4382	3.1765	14.4615	0.440
	<b>39. sag</b>	2.2972	41.6253	39.3281	36.1679	91.5205	7.138
	<b>36. backprop</b>	2.9109	3.4962	0.5853	2.6743	11.9971	0.622
<b>keras</b>	<b>28. adamax</b>	2.5783	3.1920	0.6137	2.5827	8.6394	6.082
	<b>27. adam</b>	2.8614	3.0075	0.1461	2.3994	8.8657	3.620
	<b>29. nadam</b>	5.2915	12.2499	6.9584	9.8464	27.3131	3.166
	<b>26. adagrad</b>	6.5047	6.7324	0.2277	5.3578	20.2040	52.480
	<b>25. adadelta</b>	4.4960	5.1564	0.6604	3.8748	16.8817	73.920
	<b>31. sgd</b>	3.4196	3.9121	0.4925	3.1018	14.8772	13.988
	<b>30. rmsprop</b>	6.6698	13.1877	6.5179	10.5806	32.3150	2.552
<b>AMORE</b>	<b>2. ADAPTgdwm</b>	14.8661	28.9286	14.0625	15.0901	75.7312	0.084
	<b>1. ADAPTgd</b>	12.5180	28.6849	16.1669	23.0898	63.3445	0.050
	<b>4. BATCHgdwm</b>	12.0720	12.5131	0.4411	10.6879	26.8772	1.724
	<b>3. BATCHgd</b>	12.0830	12.6864	0.6034	10.8841	27.5733	1.694
<b>minpack.lm</b>	<b>33. none</b>	2.2329	2.2329	0.0000	1.7383	6.9429	0.068
<b>ANN2</b>	<b>6. rmsprop</b>	5.5210	8.5210	3.0000	7.0532	20.1249	0.076
	<b>5. adam</b>	2.7029	9.5851	6.8822	7.9902	22.2599	0.080
	<b>7. sgd</b>	11.0994	11.9128	0.8134	10.0259	25.8843	0.080
<b>deepdive</b>	<b>16. adam</b>	20.5179	20.5179	0.0000	16.5474	47.6354	0.634
	<b>19. rmsProp</b>	23.5833	23.5833	0.0000	19.1113	57.9958	0.658
	<b>18. momentum</b>	16.2557	16.2557	0.0000	13.0878	49.2409	6.236
	<b>17. gradientDescent</b>	23.6597	23.6597	0.0000	19.1988	57.7174	6.084
<b>snnR</b>	<b>54. none</b>	11.6175	11.6175	0.0000	9.5749	25.9147	0.038
<b>elmNNRepp</b>	<b>21. extremeML</b>	17.0430	18.7830	1.7400	15.4373	50.7429	0.000
<b>ELMR</b>	<b>22. extremeML</b>	37.3986	490.0535	452.6549	426.0995	1516.4503	0.010

Note: Statistics over 10 runs.

## 1.10 Result for dataset uGauss2

Table 10: Result for uGauss2

Package	Algorithm	RMSE min	RMSE median	RMSE D51	MAE median	WAE median	Time mean
<b>nlr</b>	<b>41. none</b>	2.3327	2.9754	0.6427	2.3838	9.0619	0.118
<b>rminer</b>	<b>45. none</b>	2.3604	2.3690	0.0086	1.8630	7.5477	0.082
<b>nnet</b>	<b>42. none</b>	2.3625	3.0894	0.7269	2.4567	9.4881	0.026
<b>validann</b>	<b>56. BFGS</b>	2.3571	2.3654	0.0083	1.8640	7.5013	0.792
	<b>57. CG</b>	3.5928	6.3016	2.7088	4.7543	17.8412	34.308
	<b>58. L-BFGS-B</b>	3.0632	4.0776	1.0144	3.1853	13.2715	0.840
	<b>59. Nelder-Mead</b>	6.7221	7.5819	0.8598	5.9618	22.5839	29.606
	<b>60. SANN</b>	9.9162	14.2730	4.3568	11.6041	32.1496	0.204
<b>MachineShop</b>	<b>32. none</b>	2.5986	3.1210	0.5224	2.4847	9.3644	0.024
<b>traineR</b>	<b>55. none</b>	2.3713	2.6020	0.2307	2.0822	8.0308	0.018
<b>radiant.model</b>	<b>44. none</b>	2.5784	4.5795	2.0011	3.7498	11.3454	0.056
<b>monmlp</b>	<b>34. BFGS</b>	3.0438	4.6769	1.6331	3.6151	13.7834	0.224
	<b>35. Nelder-Mead</b>	8.2728	9.5898	1.3170	6.7817	30.1673	0.412
<b>CaDENCE</b>	<b>12. optim</b>	2.4041	3.1503	0.7462	2.3988	10.7996	2.320
	<b>14. Rprop</b>	11.7899	17.6638	5.8739	12.7742	39.4495	5.810
	<b>13. psoptim</b>	11.1566	15.4210	4.2644	12.0091	37.0326	5.752
<b>h2o</b>	<b>24. first-order</b>	2.8574	3.4454	0.5880	2.7928	11.6466	4.394
<b>EnsembleBase</b>	<b>23. none</b>	2.7901	3.8821	1.0920	3.0341	11.9878	0.030
<b>caret</b>	<b>15. none</b>	2.4240	3.2653	0.8413	2.6499	9.8629	0.086
<b>brnn</b>	<b>11. Gauss-Newton</b>	2.3781	3.5508	1.1727	2.9377	10.2283	0.048
<b>qrnn</b>	<b>43. none</b>	2.6571	3.7782	1.1211	2.6654	16.0998	0.248
<b>RSNNS</b>	<b>51. Rprop</b>	3.5732	6.3892	2.8160	4.6858	19.1326	0.316
	<b>52. SCG</b>	6.2398	6.4892	0.2494	4.8401	20.2438	0.556
	<b>53. Backpropagation</b>	3.4215	4.8320	1.4105	3.8106	15.2518	0.318
	<b>47. BackpropChunk</b>	3.2955	4.7181	1.4226	3.6073	15.4886	0.332
	<b>48. BackpropMomentum</b>	3.3532	4.8150	1.4618	3.7195	15.6644	0.352
	<b>49. BackpropWeightDecay</b>	4.5703	6.7390	2.1687	5.5508	19.8255	0.320
	<b>46. BackpropBatch</b>	12.1638	14.7088	2.5450	11.5655	30.3793	3.266
	<b>50. Quickprop</b>	24.5455	25.0662	0.5207	19.9234	51.7595	3.522
<b>automl</b>	<b>8. adam</b>	8.1454	8.6420	0.4966	6.1349	30.0687	4.952
	<b>9. RMSprop</b>	3.9723	8.4113	4.4390	5.6794	30.3739	4.436
	<b>10. trainwpso</b>	6.1384	8.6501	2.5117	6.2805	24.3330	8.702
<b>deepnet</b>	<b>20. BP</b>	3.4032	6.4830	3.0798	4.8425	17.3581	0.320
<b>neuralnet</b>	<b>38. rprop+</b>	3.6133	8.6404	5.0271	6.1594	25.2448	0.064
	<b>37. rprop-</b>	3.5904	4.4020	0.8116	3.4777	14.1560	0.080
	<b>40. slr</b>	3.5678	3.7435	0.1757	2.7163	14.9299	0.190
	<b>39. sag</b>	3.3723	8.6390	5.2667	6.1167	24.7388	1.916
	<b>36. backprop</b>	4.2479	4.3983	0.1504	3.4494	12.7619	0.866
<b>keras</b>	<b>28. adamax</b>	3.8559	4.2292	0.3733	3.3180	12.7852	6.844
	<b>27. adam</b>	3.9732	6.7909	2.8177	5.1083	18.9390	3.046
	<b>29. nadam</b>	4.2819	7.1607	2.8788	5.6164	19.4085	3.374
	<b>26. adagrad</b>	5.0447	8.8810	3.8363	6.5481	25.2992	31.862
	<b>25. adadelta</b>	3.8515	3.8918	0.0403	2.8916	15.2710	51.048
	<b>31. sgd</b>	5.1907	8.8123	3.6216	6.4800	24.8510	9.784
	<b>30. rmsprop</b>	8.1016	10.2651	2.1635	7.5917	28.4689	2.276
<b>AMORE</b>	<b>2. ADAPTgdwm</b>	4.3864	10.4646	6.0782	6.6054	38.5720	0.068
	<b>1. ADAPTgd</b>	7.4794	8.1969	0.7175	6.0059	23.9130	0.054
	<b>4. BATCHgdwm</b>	9.2190	9.4697	0.2507	6.9325	27.9676	1.578
	<b>3. BATCHgd</b>	9.1582	9.7638	0.6056	7.1783	28.5231	1.552
<b>minpack.lm</b>	<b>33. none</b>	2.9795	2.9795	0.0000	2.3890	9.0540	0.050
<b>ANN2</b>	<b>6. rmsprop</b>	3.5637	4.2355	0.6718	3.2536	15.9291	0.100
	<b>5. adam</b>	3.7310	4.0836	0.3526	3.2330	12.0798	0.086
	<b>7. sgd</b>	7.4578	8.3914	0.9336	6.3470	25.3806	0.078
<b>deepdive</b>	<b>16. adam</b>	16.8843	16.8843	0.0000	11.6361	46.6218	0.618
	<b>19. rmsProp</b>	18.3426	18.3426	0.0000	12.6957	49.5614	0.620
	<b>18. momentum</b>	28.8104	28.8104	0.0000	25.1756	67.8020	6.176
	<b>17. gradientDescent</b>	28.8118	28.8118	0.0000	25.1770	67.7823	5.934
<b>snnR</b>	<b>54. none</b>	8.8419	9.4678	0.6259	6.9147	30.1105	0.044
<b>elmNNRepp</b>	<b>21. extremeML</b>	20.3911	23.5586	3.1675	20.1337	51.5475	0.000
<b>ELMR</b>	<b>22. extremeML</b>	27.8077	31.2976	3.4899	25.2946	82.5683	0.014

Note: Statistics over 10 runs.

### 1.11 Result for dataset uGauss3

Table 11: Result for uGauss3

Package	Algorithm	RMSE min	RMSE median	RMSE D51	MAE median	WAE median	Time mean
<b>nlr</b>	<b>41. none</b>	2.2991	2.8185	0.5194	2.2078	7.5077	0.110
<b>rminer</b>	<b>45. none</b>	2.3033	2.3232	0.0199	1.8528	7.0323	0.062
<b>nnet</b>	<b>42. none</b>	2.3554	3.1706	0.8152	2.5057	9.8058	0.020
<b>validann</b>	<b>56. BFGS</b>	2.3046	2.8185	0.5139	2.2078	7.5077	0.744
	<b>57. CG</b>	2.4990	3.5798	1.0808	2.7264	11.3586	37.490
	<b>58. L-BFGS-B</b>	2.5172	3.5450	1.0278	2.7137	10.7114	0.870
	<b>59. Nelder-Mead</b>	4.9577	5.3229	0.3652	4.3142	15.0154	30.822
	<b>60. SANN</b>	6.9649	10.8474	3.8825	8.3651	26.5278	0.210
<b>MachineShop</b>	<b>32. none</b>	2.3086	3.1576	0.8490	2.4940	9.7409	0.034
<b>traineR</b>	<b>55. none</b>	2.2976	2.8669	0.5693	2.2422	7.9607	0.020
<b>radiant.model</b>	<b>44. none</b>	2.6848	3.4127	0.7279	2.7026	10.5356	0.046
<b>monmlp</b>	<b>34. BFGS</b>	2.9135	3.5782	0.6647	2.8497	10.8707	0.220
	<b>35. Nelder-Mead</b>	5.8001	7.3161	1.5160	5.7523	20.8098	0.424
<b>CaDENCE</b>	<b>12. optim</b>	2.4116	2.8622	0.4506	2.2233	7.8710	2.334
	<b>14. Rprop</b>	9.1862	21.8896	12.7034	15.2197	62.4249	5.790
	<b>13. psoptim</b>	10.8502	14.8615	4.0113	11.6021	34.5796	5.746
<b>h2o</b>	<b>24. first-order</b>	2.9525	3.2931	0.3406	2.5635	9.9032	4.706
<b>EnsembleBase</b>	<b>23. none</b>	2.3829	3.1571	0.7742	2.4537	9.9903	0.036
<b>caret</b>	<b>15. none</b>	2.4976	3.1181	0.6205	2.3743	9.6517	0.080
<b>brnn</b>	<b>11. Gauss-Newton</b>	2.8273	3.1966	0.3693	2.5109	10.0153	0.026
<b>qrnn</b>	<b>43. none</b>	2.7773	3.9015	1.1242	2.8959	13.2058	0.208
<b>RSNNS</b>	<b>51. Rprop</b>	2.9609	8.8744	5.9135	6.0200	30.1870	0.352
	<b>52. SCG</b>	3.3416	3.7196	0.3780	2.9009	11.2479	0.538
	<b>53. Backpropagation</b>	3.0294	3.2409	0.2115	2.6432	9.4181	0.316
	<b>47. BackpropChunk</b>	2.9280	3.8323	0.9043	3.0061	11.0342	0.320
	<b>48. BackpropMomentum</b>	2.9271	3.2533	0.3262	2.6246	8.8194	0.320
	<b>49. BackpropWeightDecay</b>	2.8723	3.0215	0.1492	2.3931	8.5837	0.360
	<b>46. BackpropBatch</b>	6.7104	9.7422	3.0318	7.6756	23.2263	3.274
	<b>50. Quickprop</b>	27.8595	28.8119	0.9524	22.8776	59.8157	3.518
<b>automl</b>	<b>8. adam</b>	3.1214	3.4986	0.3772	2.7250	9.7689	4.984
	<b>9. RMSprop</b>	3.5555	3.7519	0.1964	3.0224	11.8905	4.398
	<b>10. trainwpso</b>	4.8318	6.6613	1.8295	4.7306	20.0899	6.586
<b>deepnet</b>	<b>20. BP</b>	3.5001	3.7035	0.2034	2.9180	12.1143	0.300
<b>neuralnet</b>	<b>38. rprop+</b>	2.5491	3.6200	1.0709	2.8585	10.5212	0.066
	<b>37. rprop-</b>	2.6706	3.6253	0.9547	2.8533	9.9037	0.036
	<b>40. slr</b>	2.8318	3.7840	0.9522	3.0024	10.4000	0.102
	<b>39. sag</b>	2.7247	3.8181	1.0934	2.8290	11.6167	1.016
	<b>36. backprop</b>	3.8802	4.2928	0.4126	3.4361	11.4729	0.302
<b>keras</b>	<b>28. adamax</b>	2.5583	2.9976	0.4393	2.3357	9.7110	5.008
	<b>27. adam</b>	3.2791	4.0840	0.8049	3.1478	11.9057	2.296
	<b>29. nadam</b>	3.2682	3.4600	0.1918	2.8802	9.8016	2.708
	<b>26. adagrad</b>	3.4760	4.8342	1.3582	3.7910	14.2616	13.448
	<b>25. adadelta</b>	3.9293	4.3921	0.4628	3.3906	13.1185	19.964
	<b>31. sgd</b>	4.8610	5.1138	0.2528	3.9761	15.1244	5.274
	<b>30. rmsprop</b>	5.4568	6.4457	0.9889	5.4036	15.4436	1.890
<b>AMORE</b>	<b>2. ADAPTgdwm</b>	4.4658	5.1079	0.6421	3.6709	20.0320	0.080
	<b>1. ADAPTgd</b>	4.7958	4.8043	0.0085	3.9113	12.5672	0.046
	<b>4. BATCHgdwm</b>	5.0868	5.2355	0.1487	4.1127	14.7918	1.566
	<b>3. BATCHgd</b>	5.0863	5.2682	0.1819	4.1337	14.8772	1.556
<b>minpack.lm</b>	<b>33. none</b>	3.1472	3.1472	0.0000	2.4837	9.7293	0.040
<b>ANN2</b>	<b>6. rmsprop</b>	2.9727	3.2852	0.3125	2.5929	9.6420	0.082
	<b>5. adam</b>	3.1354	3.6437	0.5083	2.8900	10.5979	0.082
	<b>7. sgd</b>	4.8318	4.8821	0.0503	3.8567	14.4928	0.080
<b>deepdive</b>	<b>16. adam</b>	10.0466	10.0466	0.0000	7.3484	31.6838	0.624
	<b>19. rmsProp</b>	20.7977	20.7977	0.0000	15.4531	49.0396	0.622
	<b>18. momentum</b>	32.2413	32.2413	0.0000	27.7055	70.2077	6.146
	<b>17. gradientDescent</b>	32.2441	32.2441	0.0000	27.7063	70.1972	5.946
<b>snnR</b>	<b>54. none</b>	5.2818	5.2818	0.0000	4.0957	15.6475	0.032
<b>elmNNRepp</b>	<b>21. extremeML</b>	8.4445	19.1869	10.7424	16.8753	35.7678	0.000
<b>ELMR</b>	<b>22. extremeML</b>	31.0706	41.2586	10.1880	33.0051	112.2972	0.010

Note: Statistics over 10 runs.

## 1.12 Result for dataset uNeuroOne

Table 12: Result for uNeuroOne

Package	Algorithm	RMSE min	RMSE median	RMSE D51	MAE median	WAE median	Time mean
<b>nlr</b>	<b>41. none</b>	0.2830	0.2830	0.0000	0.2313	0.5675	0.008
<b>rminer</b>	<b>45. none</b>	0.2830	0.2830	0.0000	0.2313	0.5675	0.004
<b>nnet</b>	<b>42. none</b>	0.2830	0.2830	0.0000	0.2313	0.5675	0.000
<b>validann</b>	<b>56. BFGS</b>	0.2830	0.2830	0.0000	0.2313	0.5675	0.104
	<b>57. CG</b>	0.2830	0.2830	0.0000	0.2313	0.5675	23.762
	<b>58. L-BFGS-B</b>	0.2830	0.2830	0.0000	0.2313	0.5675	0.222
	<b>59. Nelder-Mead</b>	0.3256	0.3341	0.0085	0.2793	0.8397	9.006
	<b>60. SANN</b>	0.3084	0.3344	0.0260	0.2773	0.6937	0.168
<b>MachineShop</b>	<b>32. none</b>	0.2830	0.2830	0.0000	0.2313	0.5675	0.010
<b>traineR</b>	<b>55. none</b>	0.2830	0.2830	0.0000	0.2313	0.5675	0.000
<b>radiant.model</b>	<b>44. none</b>	0.2830	0.2830	0.0000	0.2313	0.5677	0.010
<b>monmlp</b>	<b>34. BFGS</b>	0.2831	0.2834	0.0003	0.2312	0.5810	0.194
	<b>35. Nelder-Mead</b>	0.3020	0.3266	0.0246	0.2601	0.6451	0.224
<b>CaDENCE</b>	<b>12. optim</b>	0.2831	0.2831	0.0000	0.2310	0.5816	0.298
	<b>14. Rprop</b>	0.3054	0.3248	0.0194	0.2626	0.7872	2.786
	<b>13. psoptim</b>	0.5523	0.7577	0.2054	0.5800	1.9676	4.244
<b>h2o</b>	<b>24. first-order</b>	0.2831	0.2832	0.0001	0.2331	0.5539	3.344
<b>EnsembleBase</b>	<b>23. none</b>	0.2826	0.2831	0.0005	0.2326	0.5543	0.012
<b>caret</b>	<b>15. none</b>	0.2904	0.2946	0.0042	0.2437	0.6434	0.010
<b>brnn</b>	<b>11. Gauss-Newton</b>	0.3523	0.3523	0.0000	0.2848	0.8271	0.008
<b>qrnn</b>	<b>43. none</b>	0.2939	0.2939	0.0000	0.2258	0.7231	0.094
<b>RSNNS</b>	<b>51. Rprop</b>	0.2830	0.3141	0.0311	0.2531	0.7252	0.076
	<b>52. SCG</b>	0.2855	0.6216	0.3361	0.5100	1.4782	0.104
	<b>53. Backpropagation</b>	0.2834	0.3135	0.0301	0.2457	0.7675	0.082
	<b>47. BackpropChunk</b>	0.2912	0.6365	0.3453	0.5156	1.6363	0.074
	<b>48. BackpropMomentum</b>	0.2968	0.3315	0.0347	0.2742	0.7631	0.074
	<b>49. BackpropWeightDecay</b>	0.3096	0.6423	0.3327	0.5179	1.6618	0.082
	<b>46. BackpropBatch</b>	0.6867	0.6888	0.0021	0.5629	1.6534	0.788
	<b>50. Quickprop</b>	0.5304	0.5304	0.0000	0.4235	1.2829	0.764
<b>automl</b>	<b>8. adam</b>	0.2844	0.2895	0.0051	0.2378	0.6469	1.214
	<b>9. RMSprop</b>	0.2842	0.2888	0.0046	0.2403	0.6528	1.090
	<b>10. trainwpso</b>	0.2847	0.2878	0.0031	0.2350	0.5413	4.896
<b>deepnet</b>	<b>20. BP</b>	0.2830	0.2830	0.0000	0.2314	0.5653	0.084
<b>neuralnet</b>	<b>38. rprop+</b>	0.2848	0.3165	0.0317	0.2586	0.6196	0.000
	<b>37. rprop-</b>	0.2864	0.2935	0.0071	0.2480	0.6059	0.010
	<b>40. slr</b>	0.2923	0.3203	0.0280	0.2607	0.8073	0.052
	<b>39. sag</b>	0.2893	0.3212	0.0319	0.2638	0.6316	0.056
	<b>36. backprop</b>	0.2898	0.2926	0.0028	0.2423	0.5889	0.152
<b>keras</b>	<b>28. adamax</b>	0.2841	0.2864	0.0023	0.2366	0.5789	2.502
	<b>27. adam</b>	0.2869	0.2875	0.0006	0.2340	0.5886	1.320
	<b>29. nadam</b>	0.2855	0.2896	0.0041	0.2437	0.6055	1.518
	<b>26. adagrad</b>	0.2893	0.2936	0.0043	0.2429	0.5637	13.868
	<b>25. adadelta</b>	0.2871	0.2879	0.0008	0.2377	0.5887	19.378
	<b>31. sgd</b>	0.2901	0.2922	0.0021	0.2410	0.5769	3.468
	<b>30. rmsprop</b>	0.3042	0.3629	0.0587	0.3049	0.7486	1.108
<b>AMORE</b>	<b>2. ADAPTgdwm</b>	0.2854	0.2854	0.0000	0.2285	0.6436	0.028
	<b>1. ADAPTgd</b>	0.2958	0.2965	0.0007	0.2451	0.6433	0.020
	<b>4. BATCHgdwm</b>	0.2924	0.2933	0.0009	0.2419	0.6303	1.240
	<b>3. BATCHgd</b>	0.2931	0.2935	0.0004	0.2421	0.6309	1.232
<b>minpack.lm</b>	<b>33. none</b>	1.2720	1.2720	0.0000	1.1104	2.5150	0.004
<b>ANN2</b>	<b>6. rmsprop</b>	0.2904	0.2912	0.0008	0.2376	0.6015	0.008
	<b>5. adam</b>	0.3082	0.3485	0.0403	0.2776	0.7493	0.008
	<b>7. sgd</b>	0.3069	0.3088	0.0019	0.2535	0.6226	0.010
<b>deepdive</b>	<b>16. adam</b>	0.2946	0.2946	0.0000	0.2459	0.5582	0.562
	<b>19. rmsProp</b>	0.3161	0.3161	0.0000	0.2695	0.5981	0.552
	<b>18. momentum</b>	0.3544	0.3544	0.0000	0.3001	0.6152	5.348
	<b>17. gradientDescent</b>	0.3666	0.3666	0.0000	0.3105	0.6748	5.160
<b>snnR</b>	<b>54. none</b>	0.6793	0.6793	0.0000	0.5564	1.6288	0.004
<b>elmNNRepp</b>	<b>21. extremeML</b>	0.8650	0.9526	0.0876	0.7905	2.2943	0.000
<b>ELMR</b>	<b>22. extremeML</b>	0.9735	1.0466	0.0731	0.8640	2.4817	0.000

Note: Statistics over 10 runs.



### 1.13 Score probabilities

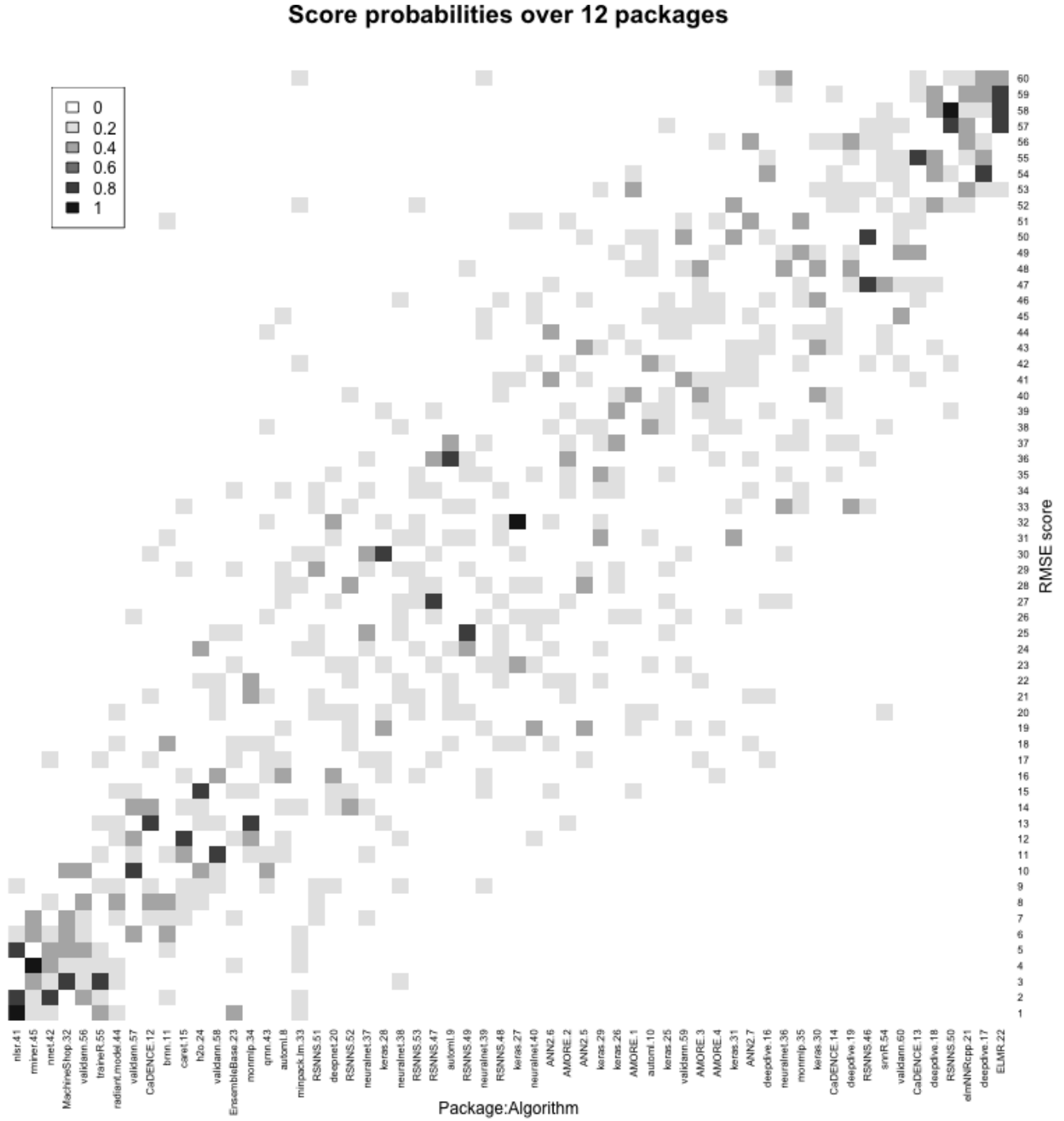


Figure 1: Score probabilities of package:algorithm

## 2 Additionnal materials on the large dataset bWoodN1

### 2.1 Summary statistics for non top-5 packages

The table provides the summary statistics of the best run of NN packages over two runs on the large dataset bWoodN1.

Table 13: Result for bWoodN1

Package	Algorithm	RMSE	MSE	MAE	WAE	time
<b>CaDENCE</b>	cadence.fit_optim_02	4.5432	20.6404	3.7038	17.0066	295.830
<b>h2o</b>	deeplearning_first-order_01	4.5720	20.9034	3.6973	17.4539	140.482
<b>monmlp</b>	monmlp.fit_BFGS_02	4.9109	24.1166	4.1281	15.4511	10.269
<b>monmlp</b>	monmlp.fit_Nelder-Mead_01	6.6516	44.2435	5.3684	24.6525	39.637
<b>radiant.model</b>	nn_none_01	10.9572	120.0608	8.7685	42.6188	0.008
<b>traineR</b>	train.nnet_none_02	4.9206	24.2125	4.1268	15.7517	3.774

*Note:* Time in seconds

## 2.2 Graphics for top-5 packages

Figures below provides some insights where a package performs reasonably well with respect to one explanatory variable and where the fit misses the correct behavior of an explanatory variable. It displays the average response per rounded explanatory variable for the predicted, the empirical and the theoretical values. That is, the empirical value and the predicted value for the  $j$ th explanatory variable are respectively computed at  $x$ -value  $x$  as

$$\bar{y}_j^{emp}(x) = \frac{1}{n_x} \sum_{i=1}^n y_i 1_{r(x_{i,j})=x}, \quad \bar{y}_j^{pred}(x) = \frac{1}{n_x} \sum_{i=1}^n \hat{y}_i 1_{r(x_{i,j})=x}, \quad n_x = \sum_{i=1}^n 1_{r(x_{i,j})=x},$$

where  $r()$  denotes the round function with two decimal places and  $y_i, \hat{y}_i$  stand respectively for the  $i$ th observed response and the  $i$ th predicted response. For instance, **MachineShop**, **nnet**, **nlr** do not correctly capture the sinusoidal aspect of explanatory variable  $x_5$  on the expected response, whereas **rminer**, **validann** miss the increasing non-linear trend of explanatory variable  $x_1$  on the expected response.

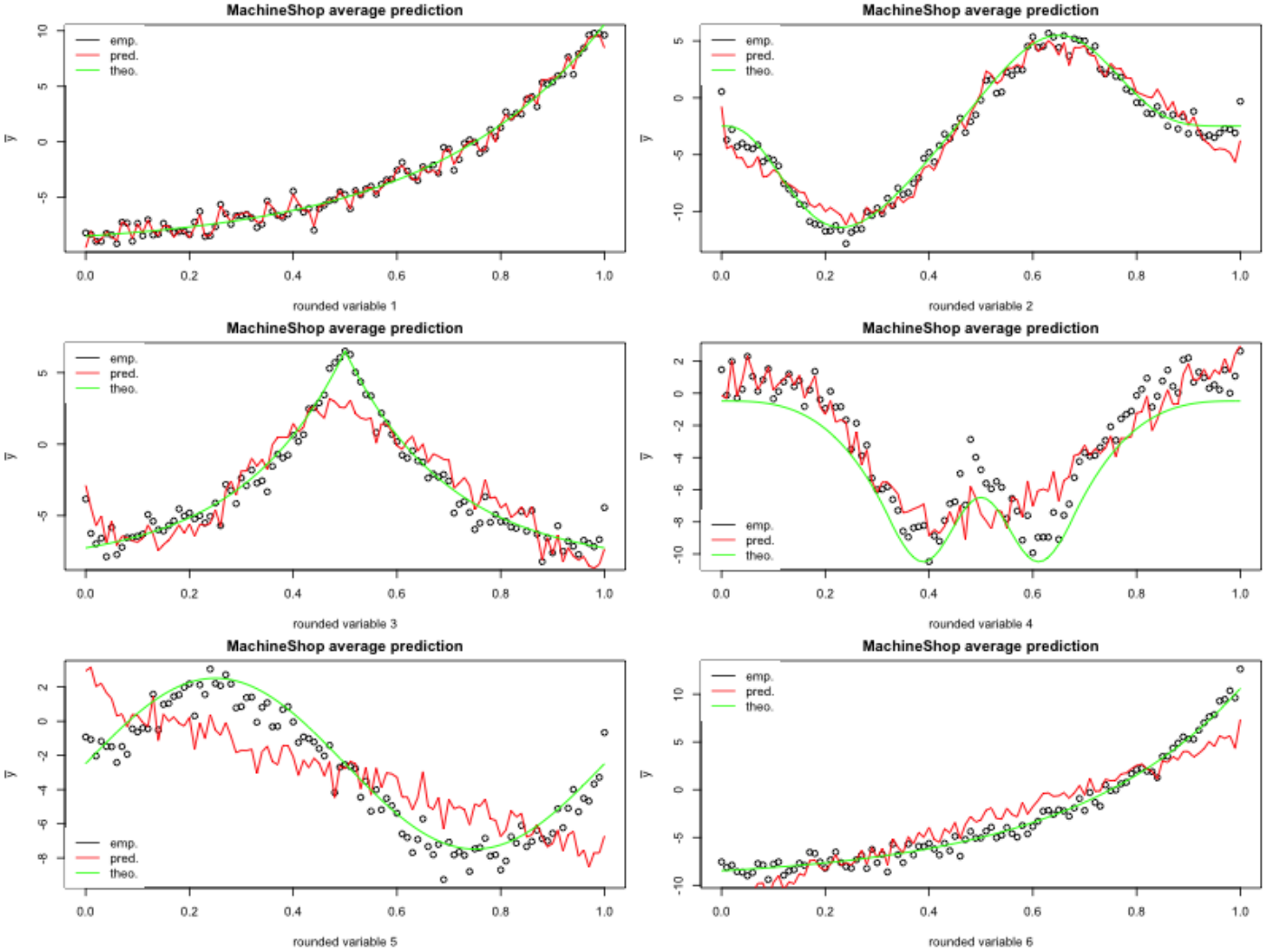


Figure 2: Average predicted mean per explanatory variable for MachineShop

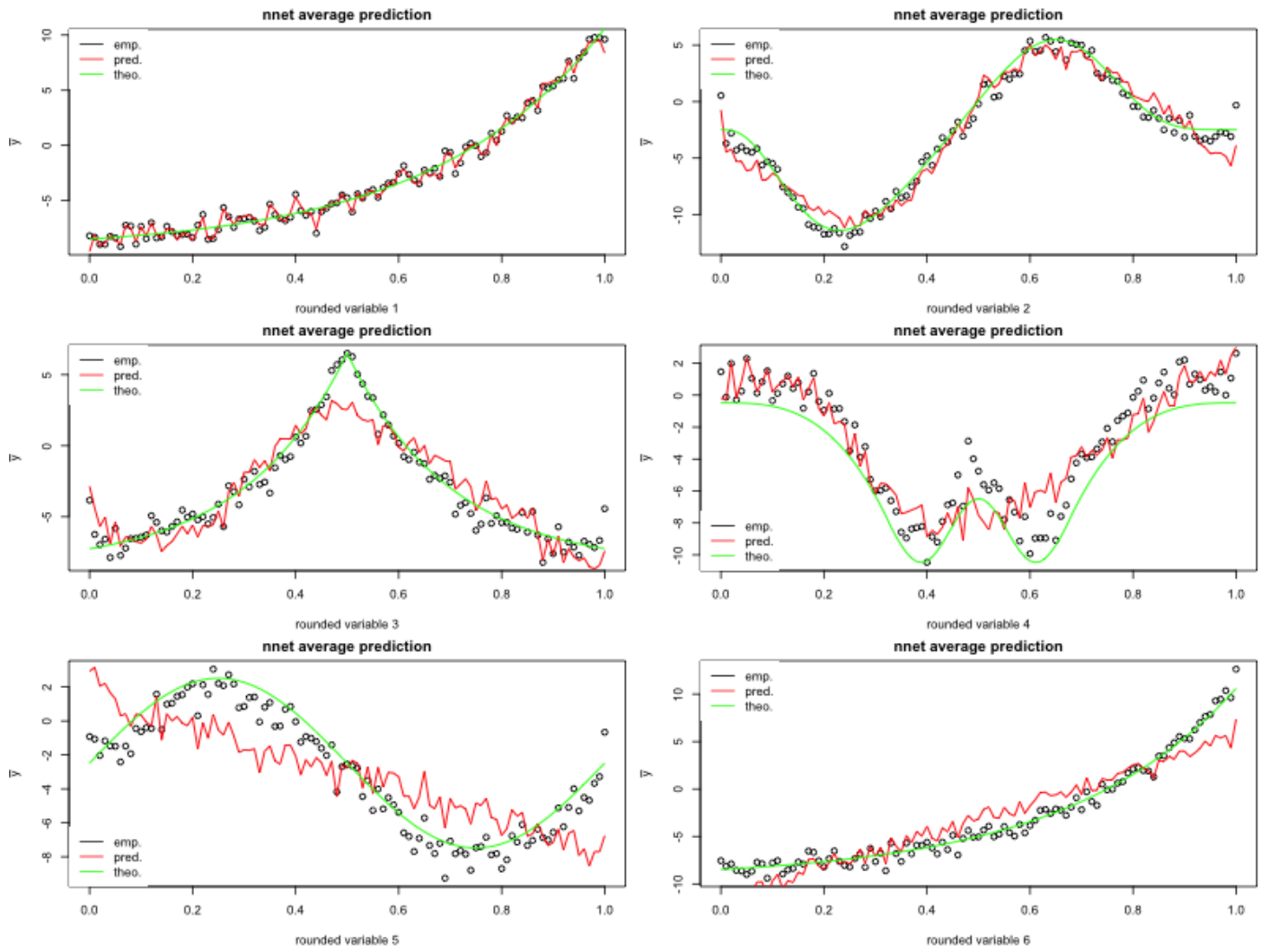


Figure 3: Average predicted mean per explanatory variable for `nnet`

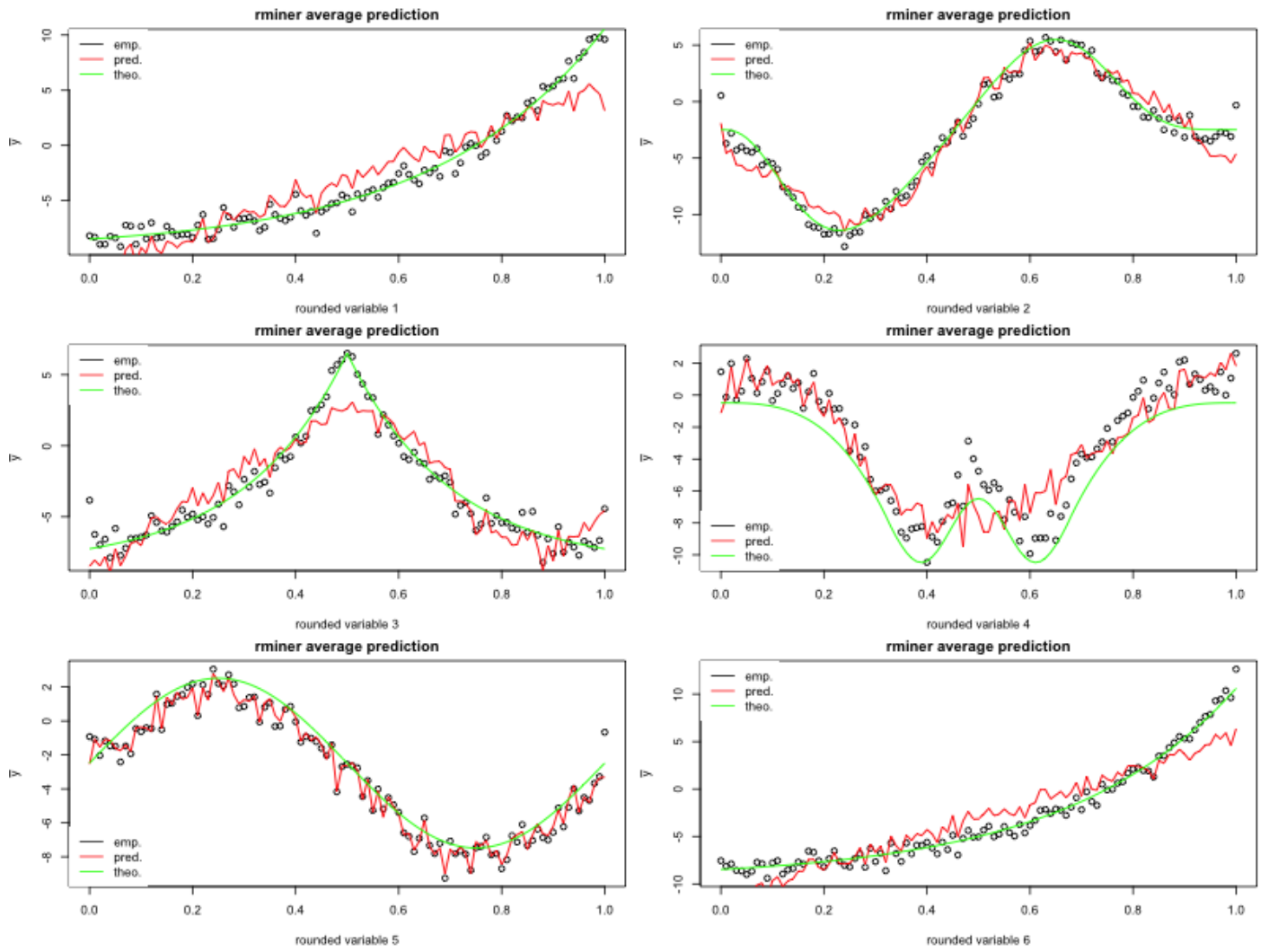


Figure 4: Average predicted mean per explanatory variable for `rminer`

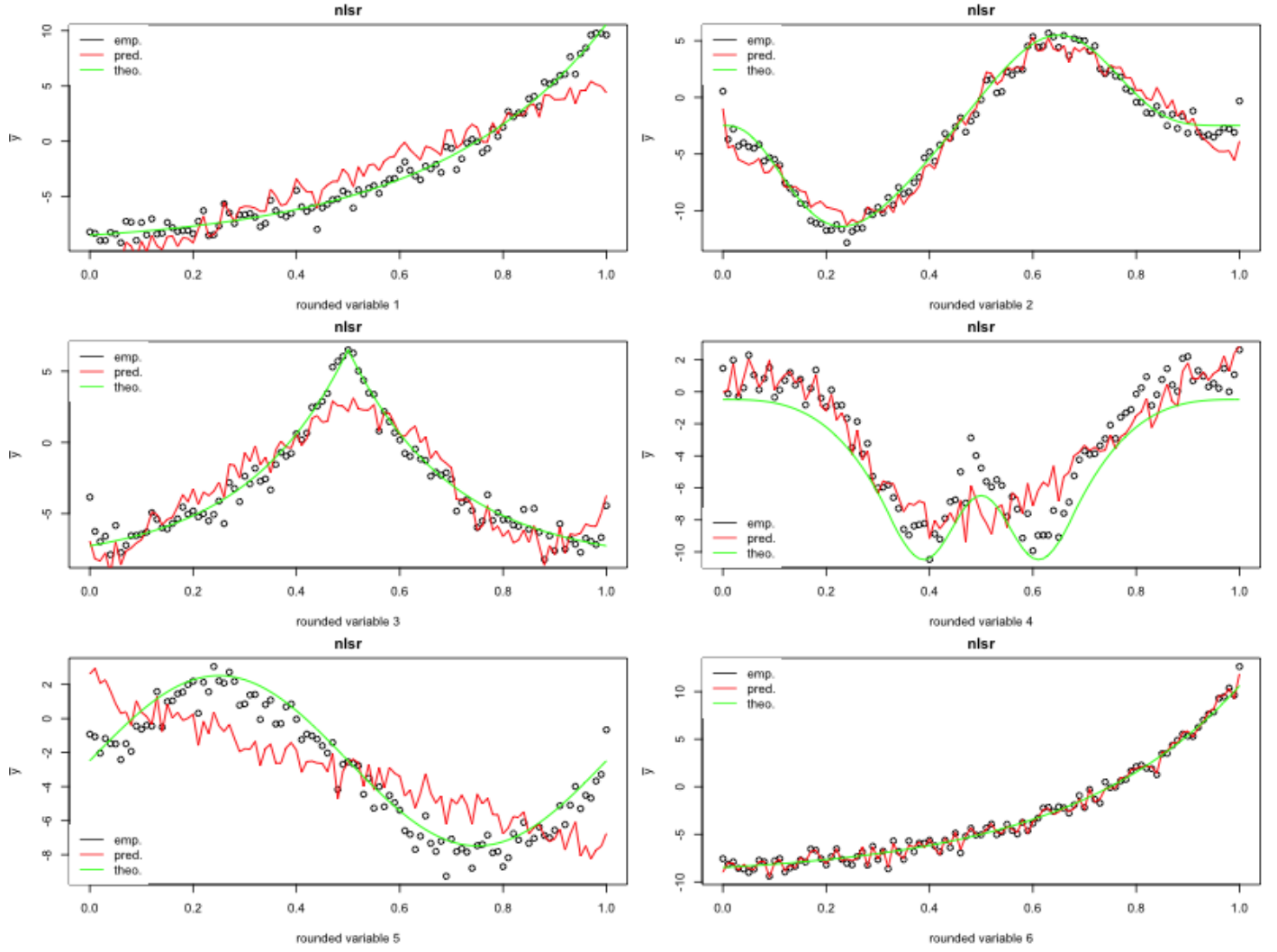


Figure 5: Average predicted mean per explanatory variable for `nlsr`

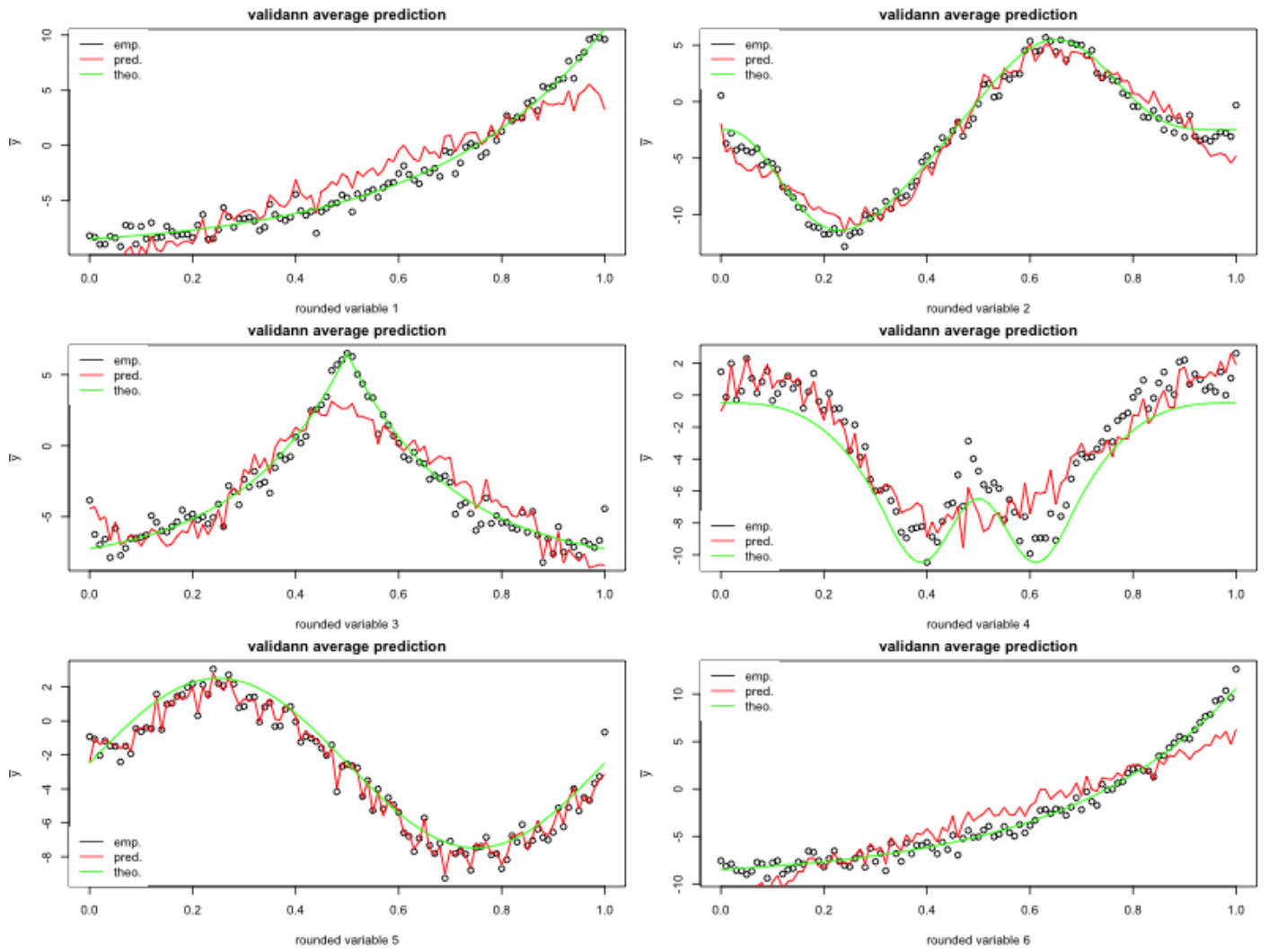


Figure 6: Average predicted mean per explanatory variable for validann