How to train Apellis

1. make your data ready for training in a csv file

endpoint (y) - properties (X)

	A	B	C	D	E	F	G	H	I	J
1	NP ID	net.c	class	lspri_synth	lspri_serur	lspri_relati	lspri_diff	lspri_rel_c	zav_synth	zav_serum
2	G15.AC	-5.1839	1	0.18253	0.454404	2.489473	0.271874	1.489473	22.36	57.53
3	G15.AHT	-1.00854	0	0.45821	0.525747	1.147394	0.067537	0.147394	30.95	90.06
4	G15.Ala-SH	-5.50439	1	0.223534	0.274761	1.22917	0.051227	0.22917	22.64	44.43
5	G15.Asn-SH	-5.67669	1	0.27362	0.327264	1.196055	0.053645	0.196055	23.09	37.75
6	G15.AUT	-1.31567	0	0.365436	0.389573	1.066051	0.024138	0.066051	23.8	55.98
7	G15.CALNN	-7.13797	1	0.20691	0.265327	1.282332	0.058417	0.282332	25.22	38.8
8	G15.CIT	-5.41982	1	0.210431	0.292836	1.391602	0.082405	0.391602	18.65	54.03
9	G15.CTAB	-5.86229	0	0.326142	0.365811	1.121632	0.039669	0.121632	15.6	59.7
10	G15.DDT@BDHDA	-7.29449	0	0.266579	0.317134	1.189643	0.050555	0.189643	23.15	47.03
11	G15.DDT@CTAB	-7.59005	0	0.275461	0.324751	1.178939	0.049291	0.178939	20.53	48.4
12	G15.DDT@DOTAP	-1.12756	0	0.276498	0.297223	1.074954	0.020725	0.074954	28.17	47.34
13	G15.DDT@ODA	-6.1218	0	0.309989	0.367331	1.184981	0.057342	0.184981	33.6	58.95
14	G15.DDT@SA	-6.8039	1	0.395907	0.320779	0.810238	-0.07513	-0.18976	82.41	59.93
15	G15.DDT@SDS	-7.67595	1	0.465011	0.359906	0.773974	-0.1051	-0.22603	27.94	100.13
16	G15.DTNB	-6.08314	1	0.241281	0.4					
17	G15.F127	-5.36112	1	0.175809	0.2					
18	G15.Gly-SH	-4.97528	1	0.261762	0.4					
19	G15.HDA	-0.27033	0	0.243839	0.2					
20	G15.LA	-5.96398	1	0.236782	0.					
21	G15.MAA	-6.14203	1	0.281334	0.4					
22	G15.MBA	-5.38142	1	0.293044	0.5					
23	G15.MES	-3.19932	1	0.226324	0.4					
24	G15.Met-SH	-5.9284	1	0.215656	0.					
25	G15.MHA	-5.73543	1	0.278335	0.3					
26	G15.MHDA	-5.77833	1	0.242059	0.2					
27	G15.MPA	-5.39593	1	0.303644	0.4					
28	G15.MSA	-6.10482	1	0.255577	0.4					

train

test

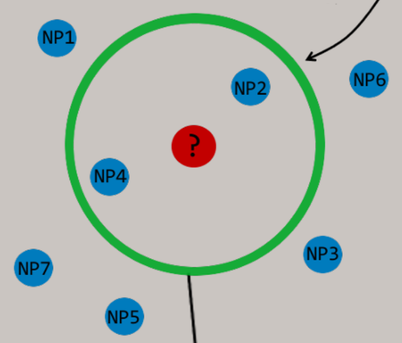
external validation

y: a property difficult to measure due to time, resources or ethical restrictions

chromosome (selected variables)

0 0 1 0 1 0 0 1 0 1 0 0 0 0 0 1 0 2.174

threshold



Apellis | HOME | NUMERICAL SINGLE CRITERION | NUMERICAL MULTIPLE CRITERIA | CLASS SINGLE CRITERION | CLASS MULTIPLE CRITERIA

READ-ACROSS TRAINING | OBTAIN PREDICTIONS

SPECIFICATIONS | PROBABILITIES

Choose training file

Use demo dataset

DEMO

Scaling of raw data

Partitioning method

Kennard-Stone

Random

Training ratio

0.66

Number of chromosomes

20

Number of generations

100

Number of training samples with a prediction

0.6

Weight of selected variables

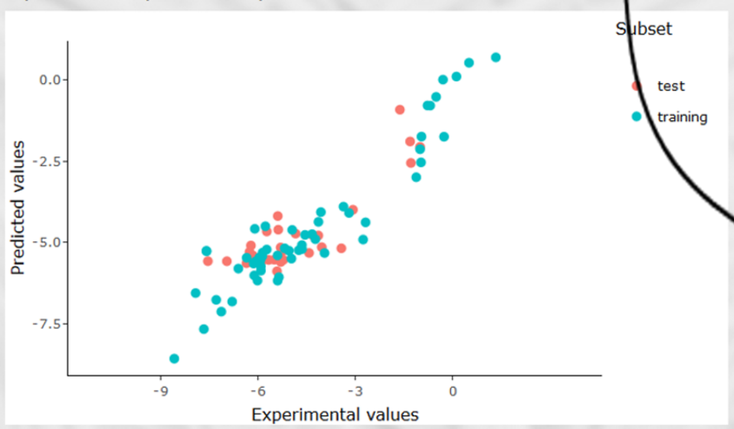
0.05

TRAIN

2. run training

3. results

Experimental vs. predicted endpoint values for the two sets



accuracy statistics

Metric	Summary
best generation	97
q2	0.731
MSE	0.805
score in test set	0.333
# samples	29
threshold	1.390
# variables	44

selected variables

Variables

lspri_serum, lspri_rel_ch, vol_synth, vol_serum, num_synth, num_serum, int_serum, hdlayer_synth, hdlayer_serum, zav_ch, int_ch, zav_rel, vol_rel, int_rel, zp_ch, zp_rel, zp_serum_mag, AS.total, P01024, P0C0L4, P02649, P10909, P01009, P04114, P00734, P0C0L5, P01008, P04196, P02656, P05154, P06396, P02654, P03952, P02760, P02655, P00748, P00740, P27169, P18065, P00450, P01019, P02671, P15169, Q13790

Experimental vs. predicted endpoint values for the test set

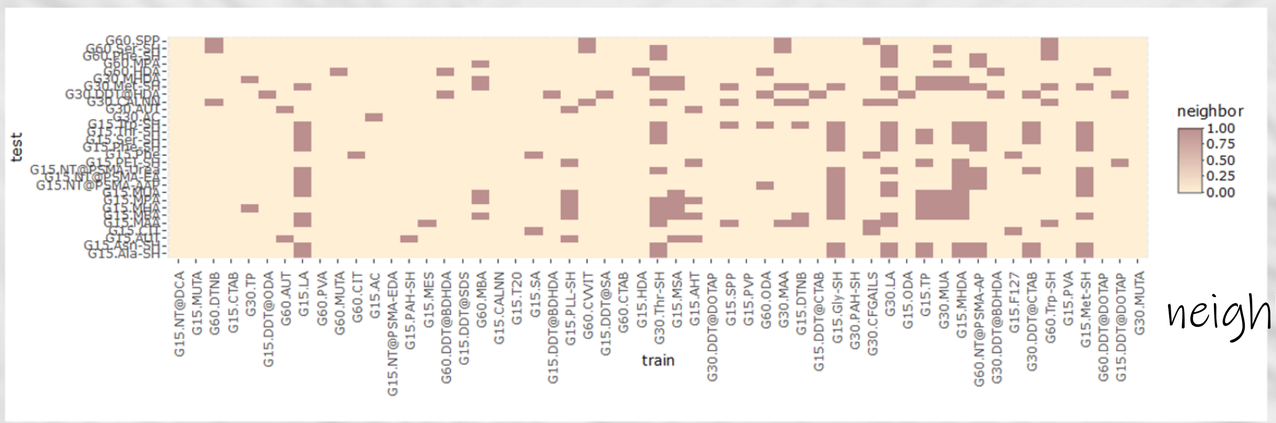
Show 10 entries | Search:

	Experimental values	Predicted values
G15.Ala-SH	-5.504	-5.536
G15.Asn-SH	-5.677	-5.542
G15.AUT	-1.316	-1.89
G15.CIT	-5.42	-5.892
G15.MAA	-6.142	-5.513
G15.MBA	-5.381	-4.609
G15.MHA	-5.735	-4.664
G15.MPA	-5.396	-4.189
G15.MUA	-4.847	-4.73
G15.NT@PSMA-AAP	-6.224	-5.097

Showing 1 to 10 of 29 entries | Previous 1 2 3 Next

predictions

Neighbors heatmap



neighbors

DOWNLOAD TRAINING RESULTS

Model title

PCF 07-01

DOWNLOAD MODEL

4. that's it! now, download your model!

Steps:

1. insert your dataset
2. press the train button
3. wait for training
4. download your model!