

Session: 2, Model: RF_T_Half_Life

Fri November 9 2012, 10:45

Modeled property: HTL

Modeling technique: Random Forests Classification

Model statistics:

	Number	Kappa	Accuracy
TRN	998	0.9975	0.999
VAL	244	0.5717	0.8402

Parameters used:

Descriptor pre-selection:

- Threshold for minimum occurrence: 4%
- Threshold for minimum standard deviation: 0.0005
- Threshold for maximum correlation between descriptors: 0.95

Descriptors remaining after pre-selection: 181

Descriptors used in the model: 181

Model details:

Data was classified into 2 classes

- $Y \leq 0.5$, class low
- $0.5 < Y$, class high

Confusion matrices

Training set

		Predicted	
		low	high
Observed	low	281	1
	high	0	716

- Sensitivity for low : 0.996
- Specificity for low : 1
- Sensitivity for high : 1
- Specificity for high : 0.999

Validation set

		Predicted	
		low	high
Observed	low	41	22
	high	17	164

- Sensitivity for low : 0.651
- Specificity for low : 0.707
- Sensitivity for high : 0.906
- Specificity for high : 0.882

Descriptor	Importance
logP	1
Vx	0.354127198
MW	0.293211162
NegativeCharge	0.0164625607
PositiveCharge	0.0446830057
Flex	0.0942217484
AromaticRings	0.0400582962
ERTLNotPSA	0.258709192
HBA-lip	0.078369014
HBA-prof	0.0479548462
HBD-lip	0.101057611
HBD-prof	0.0617345236

ACamideO-nh-nh2	0.0999561399
ACamideO-nh0	0.00716460543
ASamideO-nh-nh2	0.0116588417
ASamideO-nh0	0.00509715779
AbasicNH0	-0.00142928679
AbasicNH1	0.00222176523
CF3	0.00403835671
CH0Aa	0.00827568676
CH1Aa	0.0373496078
CH2Aa	0.0778875127
CH2hetero	0.210178137
CH2link	0.223180756
CH2long	0.0036973434
CH3Aa	0.0768998787
CH3hetero	0.0152149247
CamideNH0	0.0349992216
Ester	0.0114365825
HaloC	0.0545288436
NRB	0.155664936
RCamideO-nh0	0.0156713929
RSamideO-nh0	0.0290644616
Ramidine	0.00912789255
RbasicNH0	0.747550249
RbasicNH1	0.0119818626
SamideNH0	0.0534108803
aliphOH-t6	0.00536212232
allylic-oxyd-t10	0.016464252
aminoethanol0	0.0466096848
aminoethanol1	0.025119191
anycarbonyl	0.291684568
aromCl	0.0522502139
aromF	0.0603407808
aromO	0.00310735707
aryINHCO	0.05432209
basic-NH2	0.00728663616
branchedCnotRing	0.0398332067
ch2-lipo-t9	0.0390086994
ertl-33	0.0207155738
est-lact-latm-carbm-t7	0.00950282253
ether	0.0603423864
hetero-halo-di-n-arom	0.00837865844
hydroxyA	0.0103082052
hydroxylation-t8	0.0175387673
intraHbond5	0.0125713944
intraHbond6	0.0173679218
lipovolume	0.0698773637
nH0indole-like	0.0646026358
nHindole-like	0.00473486586
nonring-at	0.146795318
p-hetero-or-halo	0.106685378
phenol-pyr2r	0.00695641525
phenol	0.0120666055
phenolic-tautomer	0.0598435029
pyridine	0.0565649457
pyridones	0.0267398544
ring-join	0.0862687826
ring5-nH0	0.0670657977
ring5nH	0.0103724468
ringOdouble	0.0245688986
ringat	0.155734062
sp-carbons	0.00194860855
sp2-carbons	0.155354381
sulphates	0.0178249516

t-16-1	-0.000328959839
tert-amine-t11	0.0260436609
urea-thio	0.00368354446
xccn-t12	-0.000428548665
nC(sp2)	0.135836825
nC(sp3)	0.811534524
nOH	0.0121788522
nCO	0.178998724
nOS	0.0746114627
nX	0.113868967
nNprot	0.0842604637
ssCH2	0.170888662
dsCH	0.00197298778
aaCH	0.142348409
sssCH	0.0959260017
dssC	0.139246121
aasC	0.0960486755
aaaC	0.0660924539
ssssC	0.0397292599
sNH2	0.0172240436
ssNH	0.043682307
tN	0.0059049679
dsN	0.00347264693
aaN	0.0427431539
sssN	0.0426260196
sOH	0.0374165662
ssO	0.0362085328
sF	0.0548537262
nNneutral	0.048089955
NnH	0.0626387075
N4	0.128000006
NbN	0.0884015188
fg5	0.0158525947
BasicNHOR2AroRings	0.0281743109
BasicNH02AroRings	0.0210578088
BasicNH12AroRings	0.011385696
PRX-time1	0.0177717581
PRX-time-1	0.119950622
UB	0.261949718
PRX-time2	0.00556289544
HAS	0.0269918665
HAO	0.102906689
AliRingAttachment	0.14306964
C12	0.0191492047
C4	0.0225781482
C10	0.0517265946
C6	0.0288396496
C9	0.00493018562
C8	0.0139906211
C1	0.110206708
C11	0.0645677298
C2	0.0184046179
C26	0.00619990006
N7	0.0580560379
N8	0.0623383224
N2	0.0391170196
N1	0.0146260774
BasicGroup	0.04890747
AcidGroup	0.0738938078
H4	0.123218805
H2	0.0190054178
O3	0.0229774136
O11	0.0237488691

O9	0.0298317466
O10	0.0344317183
AroRingAttachment	0.153358176
HydrophobicGroup	0.0318986028
C5	0.257269859
C21	0.120610997
C22	0.123404175
C23	0.0563393049
C24	0.0344444066
S3	0.0367957689
ed70	0.0886184946
ed40	0.0475467294
ed80	0.048488833
ew70	0.00667269994
ew60	0.00417752378
ew75	0.0540514812
ew10	0.0198811069
ew100	0.0273896884
f004	0.168726847
f007	0.168254152
f015	0.0609316938
f244	0.0659841895
f245	0.100104548
f301	0.153543457
f407	0.113145307
f440	0.201368138
f441	0.122560672
f443	0.131791845
f444	0.12463811
q017	0.0950298384
q039	0.123739675
q040	0.327321827
q137	0.148031443
q155	0.219112501
q192	0.161159664
q257	0.0730105489
q300	0.153584599
q453	0.133530915
q457	0.0765142813
q458	0.116009891
q481	0.147145361
frg-8	0.0760425925
frg-26	0.0678285956