

Session: BBB_TrainingSet1, Model: BBB Shen training

Sun May 12 2013, 18:25

Training data set: BBB_TrainingSet1

Validation data set: BBB_TestSet1

Test data set: BBB_ExternalSet

Modeled property: Class

Modeling technique: Random Forests Classification

Model statistics:

	Number	Kappa	Accuracy
TRN	1093	1	1
VAL	499	0.8506	0.976
TEST	246	0.2646	0.6992

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: 172

Descriptors used in the model: 172

Model details:

Data was classified into 2 classes

- $Y \leq 0$, class low
- $0 < Y$, class high

Confusion matrices

Training set

		Predicted	
		low	high
Observed	low	261	0
	high	0	832

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

Validation set

		Predicted	
		low	high
Observed	low	38	11
	high	1	449

- Sensitivity for low : 0.776
- Specificity for low : 0.974
- Sensitivity for high : 0.998
- Specificity for high : 0.976

Test set

		Predicted	
		low	high
Observed	low	27	64
	high	10	145

- Sensitivity for low : 0.297

- Specificity for low : 0.73
- Sensitivity for high : 0.935
- Specificity for high : 0.694

Descriptor	Importance
logP	0.0649348497
Vx	0.0126359072
NegativeCharge	0.0022102131
PositiveCharge	0.0243411176
Flex	0.00758343609
AromaticRings	0.00714386953
ERTLNotPSA	0.412851393
HBA-prof	0.0441465005
HBD-lip	0.972503722
HBD-prof	0.0624650344
ACamideO-nh-nh2	0.00277147139
AbasicNH0	0.00141174311
AbasicNH1	0.00487845484
CH0Aa	0.00935207587
CH1Aa	0.00665358128
CH2Aa	0.00783612672
CH2hetero	0.00233777845
CH2link	0.00352154626
CH2long	9.03112596e-006
CH3Aa	0.0507947393
CH3hetero	0.00260812789
CamideNH0	0.0275108647
Ester	0.000494033215
HaloC	0.00664695119
Michael-accept	0.00304084341
NRB	0.0269274823
OHCHCdO	0.00136250595
Ocarbamate	-3.40707951e-009
RCamideO-nh-nh2	-3.40707951e-009
RCamideO-nh0	0.0515702441
RSR	0.00272497348
RbasicNH0	0.00824449304
aliphOH-t6	0.00632729987
allylic-oxyd-t10	-0.000451724365
aminoethanol0	0.0888820216
aminoethanol1	0.0533978939
anycarbonyl	0.0153972981
aromCl	0.00273796287
aromF	0.00262997928
basic-NH2	-3.40707951e-009
benzylicOH	0.00174758665
branchedCnotRing	-0.000542920083
carbonate-carbamate	0.000870786782
ch2-lipo-t9	0.00224983483
di-widhrow-cx4	0.00141377072
ertl-33	0.00972600468
ertl-35	0.000875926926
est-lact-latm-carbm-t7	0.00131506938
ether	0.0307135154
hydroxylation-t8	0.000945251726
intraHbond5	0.0331953838
intraHbond6	0.00345181697
ketone-t14	0.00214503938
ketones	0.00133465778
lipovolume	0.0205806773
nonring-at	0.0941743255
p-hetero-or-halo	0.000458016671
phenol	0.003590815
phenolic-tautomer	0.00342143863

pyridine	0.000435885013
ring-join	0.0102412663
ring5-nH0	0.000399223354
ring5nH	0.00224346854
ringOdouble	0.0383297764
ringat	0.0243517626
sp2-carbons	-0.00056158728
spiroC	0.0057657836
t-16-1	0.00134494272
tert-amine-t11	0.000953481474
urea-thio	0.000926925277
xccn-t12	0.00485920161
nC(sp2)	0.0270049665
nC(sp3)	0.0223903507
nOH	0.0262745302
nCO	0.0146692758
nOS	0.0121958861
nX	0.00497117266
nNprot	0.00711274752
ssCH2	0.0245553404
dsCH	0.00178596156
aaCH	0.0210981723
sssCH	0.00501382304
dssC	0.0440193973
aasC	0.0243352856
aaaC	0.0141827092
ssssC	0.0136433225
sNH2	0.00253703189
ssNH	0.0255986843
aaNH	0.000887741451
dsN	0.000447950239
aaN	0.0022337574
sssN	0.0119580934
ssO	0.0178648885
sF	-3.40707951e-009
dS	0.00185209734
aaS	-3.40707951e-009
sCl	0.00179784233
nNneutral	0.0177904516
NnH	0.0161231365
N4	-3.40707951e-009
NbN	0.0384002551
fg5	0.00452035153
CamideNH	0.00490986975
BasicNH0R2AroRings	0.0153562566
BasicNH02AroRings	0.00139646465
BasicNH12AroRings	0.00748466095
PRX-time1	0.00361192762
PRX-time-1	0.0308778528
UB	0.0316747241
HAS	0.00545748509
HAT	0.212113932
HAO	0.135584638
AliRingAttachment	0.0485610738
C12	0.00660444703
C4	0.0977723375
C10	-0.000455528585
C6	0.0183157455
C3	0.0155281462
C8	0.001361556
C1	0.0400162041
C11	0.0126248542
C2	0.000447950297

C26	-0.00134386402
N6	-0.000418653188
N7	0.0333310924
N8	-0.000434817193
N2	0.00136010721
N1	0.000898372382
BasicGroup	0.0244185533
AcidGroup	0.0188086703
H4	0.864587605
O3	0.00677326601
O11	-3.40707951e-009
O9	0.0156563465
O10	0.00130563194
AroRingAttachment	0.0643448755
HydrophobicGroup	0.0262914095
C5	-0.000424381491
C21	0.00308794808
C22	0.0051934896
C23	0.00432800921
C24	0.00277934549
S3	-3.40707951e-009
ed70	0.000895991223
ed20	0.0013328878
ed40	0.0014138564
ed80	-5.45142611e-005
ew10	0.00279516191
ew100	0.00436210306
f004	0.0216939561
f007	0.00775899785
f015	0.105181478
f244	-9.2557304e-005
f301	0.00303831045
f390	-0.00397801446
f407	0.000434810267
f440	-0.000903448847
f441	0.0202682652
f443	0.00211989041
q017	0.0654558465
q039	0.0106587447
q040	0.022412343
q137	1
q192	0.0425263122
q257	0.343765497
q300	0.161428556
q453	0.0356654078
q457	0.0134163694
q458	0.0621934496
q481	0.0369954854
frg-8	-0.00227369624
frg-26	-3.40707951e-009