# Session: P-gp inhibitor classification_2

Created: Mon October 15 2012, 22:15

Classification model session

Data set: P-gp inhibitor classification no duplicates (190 compounds)

Modeled property: Inhibitor

## Summary of model results

| | Trn | | Val | | Test | |
|---|---|---|---|---|---|---|
| Model | Kappa | Accuracy | Kappa | Accuracy | Kappa | Accuracy |
| RF Classification | 1 | 1 | 0.8571 | 0.9286 | 0.7083 | 0.8571 |
| DTModel1 | 0.9177 | 0.9627 | 0.6429 | 0.8214 | 0.7083 | 0.8571 |
| DTModel2 | 0.8837 | 0.9478 | 0.6429 | 0.8214 | 0.7083 | 0.8571 |
| DTModel3 | 0.8315 | 0.9254 | 0.6429 | 0.8214 | 0.7083 | 0.8571 |
| DTModel7 | 0.4568 | 0.7836 | 0.3571 | 0.6786 | 0.3913 | 0.7143 |
| DTModel10 | 0.3685 | 0.7537 | 0.2143 | 0.6071 | 0.4615 | 0.75 |
| DTModel13 | 0.7595 | 0.8955 | 0.5714 | 0.7857 | 0.7083 | 0.8571 |
| DTModel15 | 0.8476 | 0.9328 | 0.5714 | 0.7857 | 0.7083 | 0.8571 |
| DTModel16 | 0.8476 | 0.9328 | 0.5714 | 0.7857 | 0.7083 | 0.8571 |
| DTModel17 | 0.7436 | 0.8881 | 0.5 | 0.75 | 0.4842 | 0.75 |
| DTModel19 | 0.746 | 0.8881 | 0.5714 | 0.7857 | 0.4842 | 0.75 |
| DTModel20 | 0.7484 | 0.8881 | 0.5714 | 0.7857 | 0.4948 | 0.75 |
| GPOPT | 0.8362 | 0.9254 | 0.5 | 0.75 | 0.6237 | 0.8214 |

## Parameters used:

Set split:

- Training set size: 70%
- Validation set size: 15%
- Clustering with tanimoto level: 0.7

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: 150

logP, Vx, PositiveCharge, Flex, AromaticRings, ERTLNotPSA, HBA-prof, HBD-lip, HBD-prof, ACamideO-nh-nh2, ACamideO-nh0, AbasicNH0, AbasicNH1, CH0Aa, CH1Aa, CH2Aa, CH2hetero, CH2link, CH2long, CH3Aa, CH3hetero, CamideNH0, Ester, HaloC, NRB, RSR, RbasicNH0, aliphOH-t6, allylic-oxyd-t10, aminoethanol0, aminoethanol1, anycarbonyl, aromCl, aromO, arylNHCO, branchedCnotRing, ch2-lipo-t9, di-widhraw-cx4, ertl-33, ether, hydroxylation-t8, intraHbond5, intraHbond6, ketone-t14, ketones, lipovolume, nonring-at, ohccn-t17, p-hetero-or-halo, phenol, phenolic-tautomer, pyridine, ring-join, ring5-nH0, ringOdouble, ringat, sp-carbons, sp2-carbons, spiroC, tert-amine-t11, xccn-t12, nC(sp2), nOH, nOS, nX, nNprot, ssCH2, dsCH, aaCH, sssCH, dssC, aasC, aaaC, ssssC, ssNH, aaN, sssN, sF, sCl, nNneutral, N4, NbN, fg5, CamideNH, BasicNH0R2AroRings, BasicNH12AroRings, PRX-time1, PRX-time-1, UB, HAS, HAT, HAO, AliRingAttachment, C12, C4, C10, C6, C3, C8, C1, C11, C2, C26, N6, N7, N8, N2, BasicGroup, H2, O3, O9, O10, AroRingAttachment, C25, HydrophobicGroup, C5, C21, C22, C23, C24, S3, ed70, ed20, ed80, ew10, ew100, f004, f007, f015, f244, f245, f301, f390, f407, f440, f441, q017, q039, q040, q137, q139, q192, q257, q300, q453, q457, q458, q481, frg-8, frg-26

## Compound IDs for split sets for session P-gp inhibitor classification_2

Training (134 compounds)

StarDropID 2, StarDropID 4, StarDropID 5, StarDropID 6, StarDropID 7, StarDropID 9, StarDropID 11, StarDropID 12, StarDropID 13, StarDropID 14, StarDropID 16, StarDropID 17, StarDropID 22, StarDropID 24, StarDropID 25, StarDropID 26, StarDropID 27, StarDropID 28, StarDropID 29, StarDropID 32, StarDropID 36, StarDropID 37, StarDropID 39, StarDropID 41, StarDropID 42, StarDropID 43, StarDropID 44, StarDropID 48, StarDropID 50, StarDropID 51, StarDropID 52, StarDropID 53, StarDropID 54, StarDropID 55, StarDropID 57, StarDropID 60, StarDropID 61, StarDropID 62, StarDropID 64, StarDropID 66, StarDropID 67, StarDropID 68, StarDropID 69, StarDropID 70, StarDropID 72, StarDropID 73, StarDropID 75, StarDropID 76, StarDropID 78, StarDropID 79, StarDropID 80, StarDropID 81, StarDropID 82, StarDropID 83, StarDropID 85, StarDropID 87, StarDropID 88, StarDropID 89, StarDropID 91, StarDropID 95, StarDropID 97, StarDropID 99, StarDropID 102, StarDropID 103, StarDropID 107, StarDropID 108, StarDropID 109, StarDropID 111, StarDropID 112, StarDropID 114, StarDropID 115, StarDropID 117, StarDropID 120, StarDropID 121, StarDropID 122, StarDropID 123, StarDropID 124, StarDropID 125, StarDropID 131, StarDropID 133, StarDropID 134, StarDropID 135, StarDropID 137, StarDropID 138, StarDropID 139, StarDropID 140, StarDropID 141, StarDropID

142, StarDropID 143, StarDropID 144, StarDropID 145, StarDropID 147, StarDropID 148, StarDropID 149, StarDropID 150, StarDropID 151, StarDropID 152, StarDropID 154, StarDropID 155, StarDropID 156, StarDropID 157, StarDropID 158, StarDropID 160, StarDropID 161, StarDropID 162, StarDropID 163, StarDropID 164, StarDropID 165, StarDropID 166, StarDropID 168, StarDropID 169, StarDropID 171, StarDropID 172, StarDropID 173, StarDropID 174, StarDropID 175, StarDropID 176, StarDropID 177, StarDropID 178, StarDropID 182, StarDropID 184, StarDropID 185, StarDropID 188, StarDropID 189, StarDropID 190, StarDropID 192, StarDropID 194, StarDropID 195, StarDropID 198, StarDropID 199, StarDropID 200, StarDropID 201, StarDropID 202, StarDropID 203

Validation (28 compounds)

StarDropID 1, StarDropID 10, StarDropID 18, StarDropID 19, StarDropID 31, StarDropID 33, StarDropID 35, StarDropID 38, StarDropID 45, StarDropID 46, StarDropID 56, StarDropID 59, StarDropID 65, StarDropID 71, StarDropID 93, StarDropID 96, StarDropID 98, StarDropID 101, StarDropID 105, StarDropID 106, StarDropID 113, StarDropID 116, StarDropID 118, StarDropID 146, StarDropID 159, StarDropID 167, StarDropID 170, StarDropID 193

Test (28 compounds)

StarDropID 3, StarDropID 15, StarDropID 20, StarDropID 21, StarDropID 23, StarDropID 30, StarDropID 34, StarDropID 49, StarDropID 58, StarDropID 63, StarDropID 74, StarDropID 77, StarDropID 84, StarDropID 86, StarDropID 90, StarDropID 92, StarDropID 100, StarDropID 104, StarDropID 110, StarDropID 119, StarDropID 136, StarDropID 153, StarDropID 179, StarDropID 180, StarDropID 181, StarDropID 183, StarDropID 196, StarDropID 197