

# Drugs with Lasa

## Isolating the Drugs

```
In[524]:= Drugs = EntityList[drugs CHEMICALS];  
  
In[525]:= missingACs2 = Select[Drugs, MissingQ[#[ "AtomCount"]]&];  
  
In[526]:= drugsWithAc = Complement[Drugs, missingACs2];  
  
In[527]:= Length[drugsWithAc]  
Out[527]= 4475  
  
In[528]:= missingBps2 = Select[drugsWithAc, MissingQ[#[ "BoilingPoint"]]&];  
  
In[529]:= drugsWithBP = Complement[drugsWithAc, missingBps2];  
  
In[530]:= Length[drugsWithBP]  
Out[530]= 682  
  
In[532]:= drugMolecules = ParallelMap[Molecule, drugsWithBP];  
(kernel 4) MoleculeModify::atmlst : {4, 7, 9, 1, 10, 12, 15, 17, 6, 5, 8, 3, 2, 14, 11, 13, 18, 16, 24, 23, 26, 22, 20, 21, 25, 27, 28, 19, 31, 30} is  
not a valid list of atom indices or names.  
(kernel 4) Molecule::stereo : Invalid stereochemistry specification <<1>>.
```

## Descriptors

```
In[533]:= SmilesDrugs = ParallelMap[#[ "SMILES"]&, drugsWithBP];  
  
In[534]:= BpDrugs = ParallelMap[#[ "BoilingPoint"]&, drugsWithBP];  
  
In[535]:= acDrugs = ParallelMap[#[ "AtomCount"]&, drugsWithBP];  
  
In[536]:= mmDrugs = ParallelMap[MoleculeValue[#, "MolecularMass"]&, drugMolecules];  
  
In[537]:= rgDrugs = ParallelMap[MoleculeValue[#, "RadiusOfGyration"]&, drugMolecules];  
(kernel 4) Molecule::discon : Warning: coordinate generation for disconnected structures is experimental.  
(kernel 1) Molecule::discon : Warning: coordinate generation for disconnected structures is experimental.  
  
In[538]:= cnDrugs = ParallelMap[MoleculeValue[#, "HeteroatomCount"]&, drugMolecules];  
  
In[539]:= KhasDrugs = ParallelMap[MoleculeValue[#, "KierHallAlphaShape"]&, drugMolecules];  
  
In[540]:= LaSaDrugs =  
ParallelMap[MoleculeValue[#, "LabuteApproximateSurfaceArea"]&, drugMolecules];
```

```
In[541]:= SasDrugs =
  ParallelMap[MoleculeValue[#, "SyntheticAccessibilityScore"] &, drugMolecules];

In[542]:= PobfDrugs =
  ParallelMap[MoleculeValue[#, "PlaneOfBestFitDistance"] &, drugMolecules];
(kernel 4) Molecule::discon : Warning: coordinate generation for disconnected structures is experimental.
(kernel 1) Molecule::discon : Warning: coordinate generation for disconnected structures is experimental.
```

---

## Making a dataset

```
In[543]:= measurementsDrugs =
  Transpose[{SmilesDrugs, QuantityMagnitude[BpDrugs], QuantityMagnitude[acDrugs],
  QuantityMagnitude[mmDrugs], QuantityMagnitude[rgDrugs], QuantityMagnitude[
  cnDrugs], QuantityMagnitude[KhasDrugs], QuantityMagnitude[LaSaDrugs],
  QuantityMagnitude[SasDrugs], QuantityMagnitude[PobfDrugs]}];

In[544]:= variableNamesDrugs =
 {"Smiles", "Bp", "Ac", "Mm", "Rg", "Cn", "Khas", "LaSa", "Sas", "Pobf"};
```

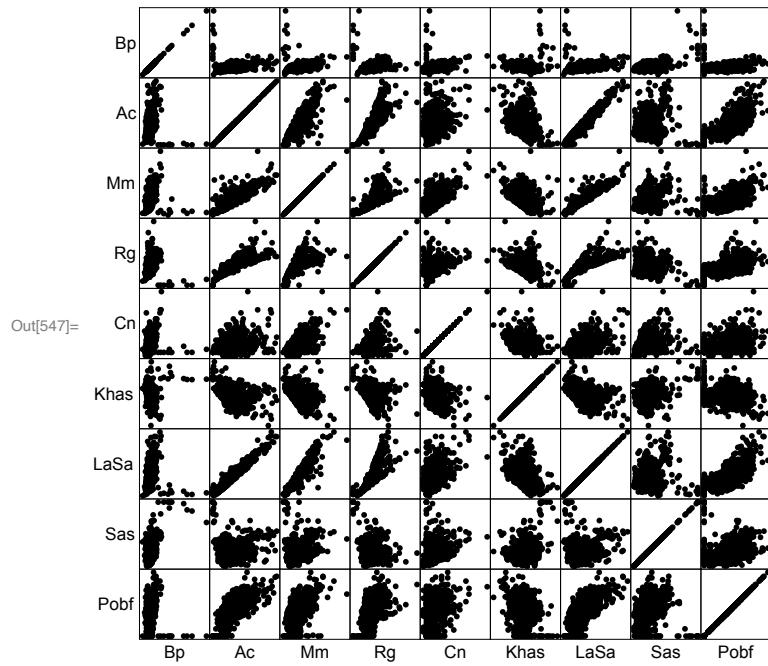
```
In[545]:= drugData =
ResourceFunction["DatasetWithHeaders"] [measurementsDrugs, variableNamesDrugs]
```

Smiles	Bp	Ac	Mm
C1=CC2=C(C3=C(C=CC=N3)C=C2)N=C1	360.	22	180.21
C(C(F)(F)F)(C(F)(F)F)(O)O	184.	13	184.03
C1=C(C(=CC(=C1C(=O)O)C(=O)O)C(=O)O)C(=O)O	516.	24	254.15
C1=NC=NN1	260.0	8	69.067
C(CCCCCBr)CCCCCO	155.0	39	265.23
C(CCl)Cl	83.0	8	98.96
CC(CCl)Cl	95.5	11	112.98
COCCOC	85.0	16	90.122
CC(CO)O	187.0	13	76.095
C1=C(C=C(C=C1Cl)Cl)Cl	208.0	12	181.44
CC1=C(C(CCC1)(C)C)C=CC(=CC=CC(=CC(=O)O)C)C	412.	50	300.44
C1=CC(=CC(=C1)[N+](=O)[O-])[N+](=O)[O-]	297.0	16	168.10
C(CO)CO	214.0	13	76.095
C(CCO)CO	230.	16	90.122
C1COCCO1	101.0	14	88.106
C(C(C(S)O)O)S	127.625	18	154.25
C(CCO)CCO	250.0	22	118.17
C(CCCN)CCCN	225.5	30	144.26
C1C2CC3CC1CC(C2)(C3)N	214.	28	151.25
C1=CC=C(C=C1)CN2C=CN=C2	310.0	22	158.20

↖ ⌘ rows 1–20 of 682 ⌘ ↘

```
In[546]:= dataFitting2 = Values[Normal[
drugData[All, {"Bp", "Ac", "Mm", "Rg", "Cn", "Khas", "LaSa", "Sas", "Pobf"}]]];
```

```
In[547]:= Graph2 = ResourceFunction["PairwiseScatterPlot"] [dataFitting2,
  "DataLabels" → {"Bp", "Ac", "Mm", "Rg", "Cn", "Khas", "LaSa", "Sas", "Pobf"}]
```



## Pearson Correlation

```
In[548]:= PearsonCorrelationTest[BpDrugs, LaSaDrugs, "TestStatistic"]
```

```
Out[548]= 0.236214
```

```
In[549]:= PearsonCorrelationTest[acDrugs, LaSaDrugs, "TestStatistic"]
```

```
Out[549]= 0.973624
```

```
In[550]:= PearsonCorrelationTest[mmDrugs, LaSaDrugs, "TestStatistic"]
```

```
Out[550]= 0.929536
```

```
In[551]:= PearsonCorrelationTest[rgDrugs, LaSaDrugs, "TestStatistic"]
```

```
Out[551]= 0.843196
```

```
In[552]:= PearsonCorrelationTest[cnDrugs, LaSaDrugs, "TestStatistic"]
```

```
Out[552]= 0.408882
```

```
In[553]:= PearsonCorrelationTest[KhasDrugs, LaSaDrugs, "TestStatistic"]
```

```
Out[553]= -0.47856
```

```
In[554]:= PearsonCorrelationTest[SasDrugs, LaSaDrugs, "TestStatistic"]
```

```
Out[554]= 0.177282
```

```
In[555]:= PearsonCorrelationTest[PobfDrugs, LaSaDrugs, "TestStatistic"]
Out[555]= 0.691775
```

## Machine Learning Functions

```
In[771]:= predictorVars2 = Normal@Values@drugData[All, {"Smiles", "Ac", "Mm", "Rg", "Pobf"}];
In[772]:= outputValue2 = Normal@drugData[All, "LaSa"];
In[773]:= fullDataSetPredict2 = MapThread[Rule, {predictorVars2, outputValue2}];
In[774]:= fullTrainingSet2 =
  RandomSample[fullDataSetPredict2, Round[0.8 * Length[fullDataSetPredict2]]];
In[775]:= fullTestingSet2 = Complement[fullDataSetPredict2, fullTrainingSet2];
In[776]:= trainingSet2Values = Rest /@ Keys[fullTrainingSet2];
In[777]:= trainingSet2Output = Values[fullTrainingSet2];
In[778]:= trainingSet2 = MapThread[Rule, {trainingSet2Values, trainingSet2Output}];
In[779]:= testingSet2Values = Rest /@ Keys[fullTestingSet2];
In[780]:= testingSet2Output = Values[fullTestingSet2];
In[781]:= testingSet2 = MapThread[Rule, {testingSet2Values, testingSet2Output}];
```

### Linear Regression

```
In[782]:= pred1p = Predict[trainingSet2, Method → "LinearRegression"]
Out[782]= PredictorFunction[ +  Input type: Mixed (number: 4)
Method: LinearRegression]
```

```
In[783]:= pm1p = PredictorMeasurements[pred1p, testingSet2]
Out[783]= PredictorMeasurementsObject[ +  Predictor: LinearRegression
Number of test examples: 135]
```

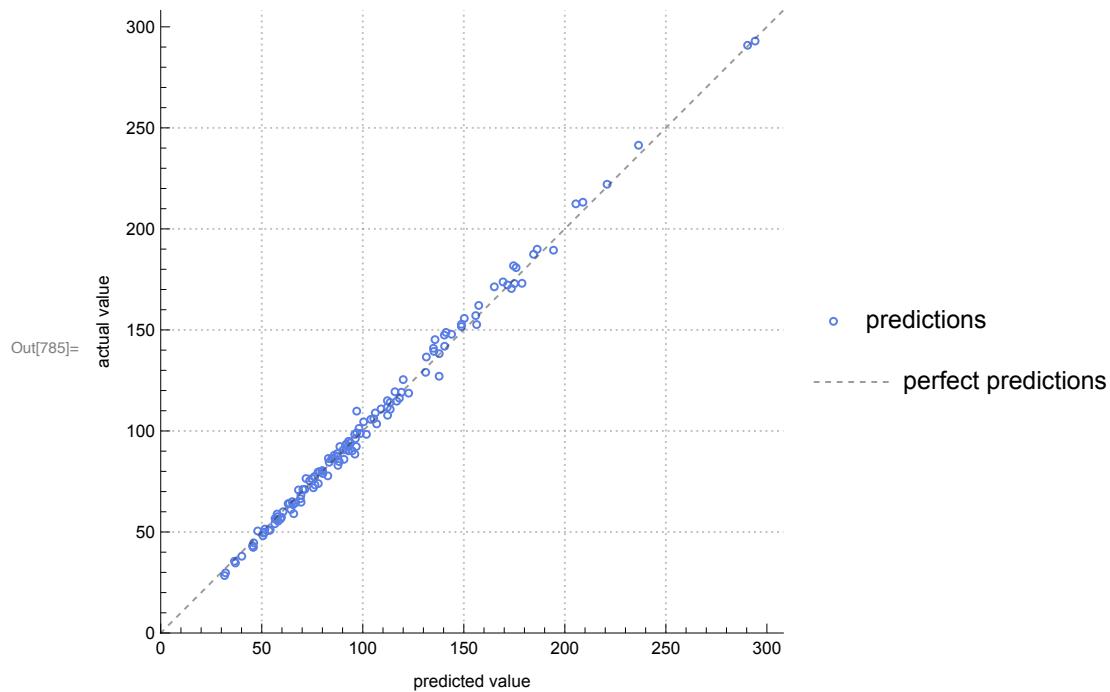
  

```
In[784]:= Dataset[AssociationMap[
  pm1p[#, ComputeUncertainty → True] &, {"StandardDeviation", "RSquared"}]]
Out[784]= 

|                   |                     |
|-------------------|---------------------|
| StandardDeviation | $3.55 \pm 0.26$     |
| RSquared          | $0.9950 \pm 0.0011$ |


```

In[785]:= PredictorMeasurements[pred1p, testingSet2, "ComparisonPlot"]



In[786]:= PredictorMeasurements[pred1p, testingSet2, "BestPredictedExamples"]

```
Out[786]= { {29, 163.264, 2.60939, 0.807387} → 98.7485,
{33, 223.272, 2.86424, 0.942411} → 119.146,
{14, 121.161, 1.99025, 0.504654} → 56.6383, {16, 142.118, 2.07814, 0.563632} → 65.007,
{25, 150.221, 2.41239, 0.571075} → 87.4417, {23, 135.21, 2.35493, 0.508427} → 80.2468,
{15, 83.134, 2.15998, 0.474992} → 51.4463, {24, 214.246, 2.65289, 0.942891} → 96.1972,
{24, 199.181, 2.80937, 0.628633} → 94.0922,
{15, 109.128, 1.86124, 0.0000194784} → 57.6031}
```

In[787]:= PredictorMeasurements[pred1p, testingSet2, "WorstPredictedExamples"]

```
Out[787]= { {18, 290.832, 2.74484, 0.762268} → 109.808,
{24, 432.983, 3.29147, 0.690096} → 127.074,
{34, 300.745, 3.32588, 0.92001} → 145.208, {37, 293.322, 3.46292, 0.965568} → 148.771,
{26, 182.156, 2.1638, 0.649499} → 88.5288, {46, 360.837, 4.40847, 0.818306} → 181.766,
{35, 313.288, 3.30954, 0.994666} → 147.437, {56, 399.629, 4.00636, 1.11738} → 212.396,
{13, 184.035, 1.94439, 0.710411} → 59.0341, {45, 324.399, 3.93421, 1.17359} → 171.338}
```

## Neural Network

In[788]:= pred2p = Predict[trainingSet2, Method → "NeuralNetwork"]

Out[788]= PredictorFunction[ Input type: Mixed (number: 4)  
Method: NeuralNetwork]

```
In[789]:= pm2p = PredictorMeasurements[pred2p, testingSet2]
```

Out[789]= PredictorMeasurementsObject[  Predictor: NeuralNetwork  
Number of test examples: 135 ]

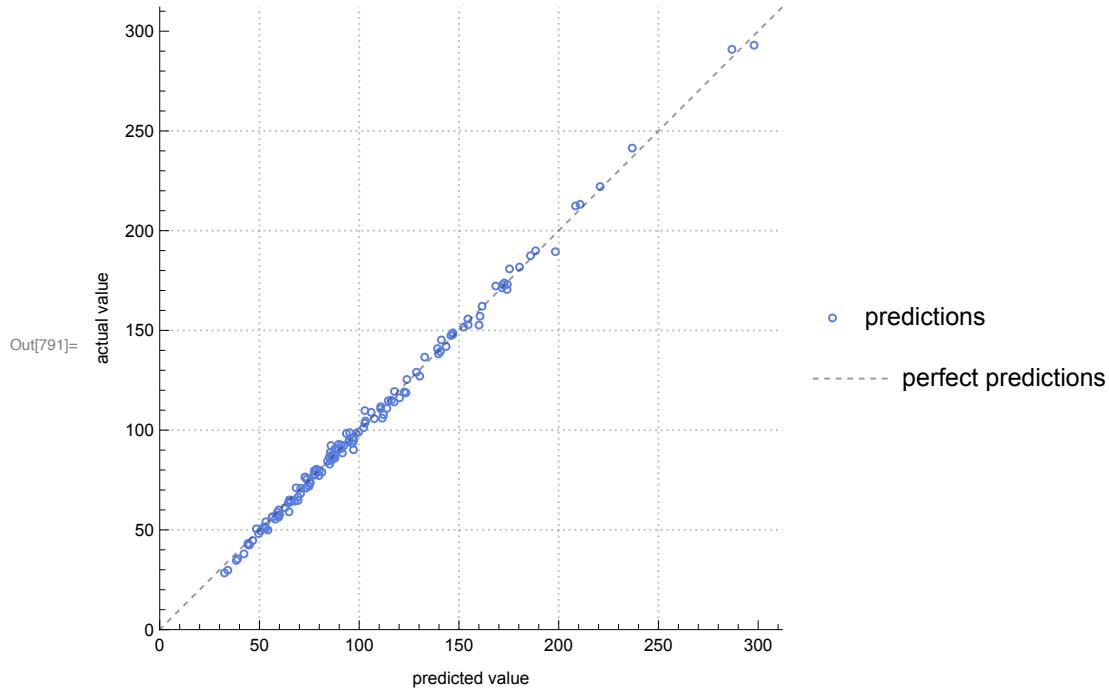
```
In[790]:= Dataset[AssociationMap[
```

```
pm2p[#, ComputeUncertainty → True] &, {"StandardDeviation", "RSquared"}]]
```

Out[790]=

StandardDeviation	$2.83 \pm 0.18$
RSquared	$0.9968 \pm 0.0007$

```
In[791]:= PredictorMeasurements[pred2p, testingSet2, "ComparisonPlot"]
```



```
In[792]:= PredictorMeasurements[pred2p, testingSet2, "BestPredictedExamples"]
```

Out[792]= { {16, 142.118, 2.07814, 0.563632} → 65.007, {30, 144.262, 3.76859, 0.543434} → 92.3053, {31, 156.269, 2.96814, 0.790133} → 98.3144, {21, 150.177, 2.45114, 0.249324} → 79.8808, {34, 179.307, 2.20948, 1.05601} → 110.697, {14, 121.161, 1.99025, 0.504654} → 56.6383, {17, 141.601, 2.16121, 0.107677} → 70.7954, {14, 124.23, 1.88208, 0.595448} → 56.2063, {14, 123.115, 2.07701, 0.140488} → 58.878, {27, 152.237, 2.01695, 1.15164} → 90.6531 }

```
In[793]:= PredictorMeasurements[pred2p, testingSet2, "WorstPredictedExamples"]
Out[793]= {{51, 416.343, 6.95058, 0.734875} → 189.457,
{38, 363.223, 4.10357, 0.871585} → 152.656,
{18, 290.832, 2.74484, 0.762268} → 109.808,
{27, 156.225, 2.84764, 0.552696} → 90.1749,
{25, 159.232, 2.46414, 0.827593} → 92.2798,
{13, 184.035, 1.94439, 0.710411} → 59.0341,
{26, 240.306, 3.31979, 0.870437} → 106.009, {50, 326.436, 5.50825, 1.02686} → 180.832,
{77, 610.565, 4.84004, 1.20624} → 292.969, {26, 183.682, 2.44856, 0.680697} → 98.32}
```

## Gradient Boosted Trees

```
In[794]:= pred3p = Predict[trainingSet2, Method → "GradientBoostedTrees"]
```

Out[794]= PredictorFunction[   Input type: Mixed (number: 4)  
Method: GradientBoostedTrees ]

```
In[795]:= pm3p = PredictorMeasurements[pred3p, testingSet2]
```

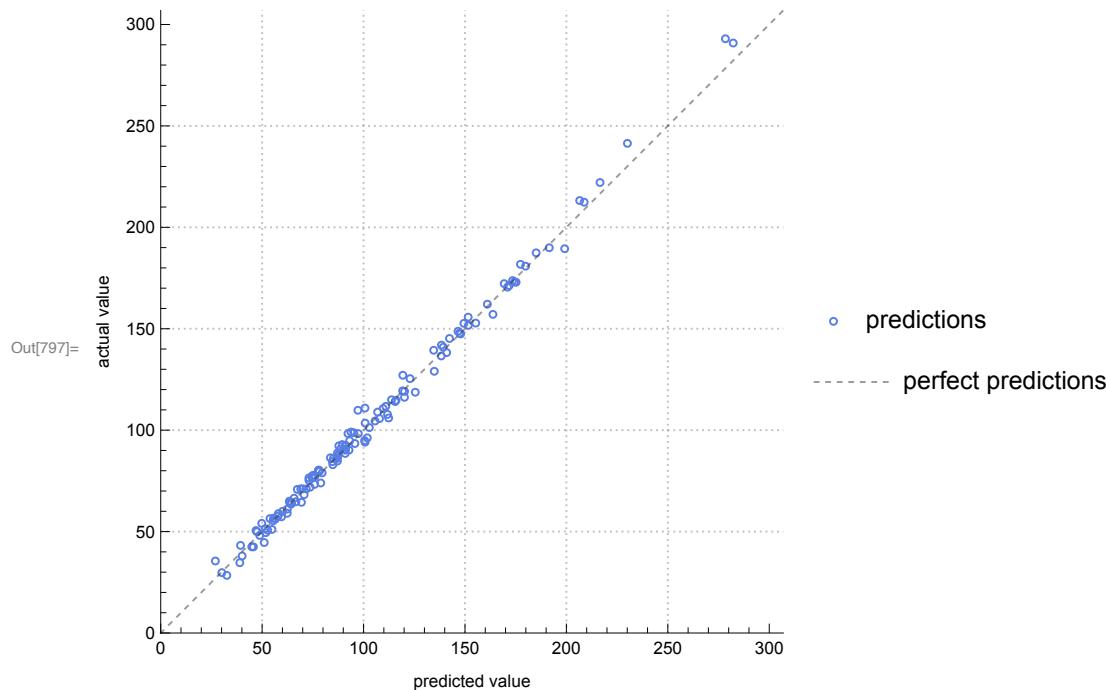
Out[795]= PredictorMeasurementsObject[   Predictor: GradientBoostedTrees  
Number of test examples: 135 ]

```
In[796]:= Dataset[AssociationMap[
pm3p[#, ComputeUncertainty → True] &, {"StandardDeviation", "RSquared"}]]
```

Out[796]= 

StandardDeviation	$3.83 \pm 0.33$
RSquared	$0.9941 \pm 0.0014$

In[797]:= PredictorMeasurements[pred3p, testingSet2, "ComparisonPlot"]



In[798]:= PredictorMeasurements[pred3p, testingSet2, "BestPredictedExamples"]

```
Out[798]= {{42, 270.376, 3.35943, 0.931283} → 151.681,
{15, 83.134, 2.15998, 0.474992} → 51.4463, {16, 103.121, 2.09829, 0.334789} → 55.3618,
{15, 130.099, 2.24907, 0.416276} → 60.0241,
{31, 230.263, 3.43024, 0.52787} → 119.398, {22, 166.176, 2.56261, 0.542645} → 84.4993,
{15, 109.128, 1.86124, 0.0000194784} → 57.6031,
{48, 307.437, 3.6929, 1.07187} → 173.759, {22, 179.175, 2.962, 0.314002} → 87.9828,
{14, 124.23, 1.88208, 0.595448} → 56.2063}
```

In[799]:= PredictorMeasurements[pred3p, testingSet2, "WorstPredictedExamples"]

```
Out[799]= {{77, 610.565, 4.84004, 1.20624} → 292.969, {18, 290.832, 2.74484, 0.762268} → 109.808,
{67, 429.604, 4.06698, 1.42492} → 241.398, {22, 305.164, 3.05661, 0.650049} → 110.901,
{51, 416.343, 6.95058, 0.734875} → 189.457, {83, 516.857, 4.37087, 1.49007} → 290.869,
{9, 68.079, 1.24751, 9.71706 × 10-6} → 35.5066,
{24, 432.983, 3.29147, 0.690096} → 127.074,
{33, 238.305, 2.84447, 0.584026} → 118.703, {41, 319.336, 3.82749, 0.621548} → 157.09}
```

## Random Forest

In[800]:= pred4p = Predict[trainingSet2, Method → "RandomForest"]

Out[800]= PredictorFunction[ Input type: Mixed (number: 4) Method: RandomForest]

```
In[801]:= pm4p = PredictorMeasurements[pred4p, testingSet2]
```

Out[801]= PredictorMeasurementsObject[  Predictor: RandomForest  
Number of test examples: 135 ]

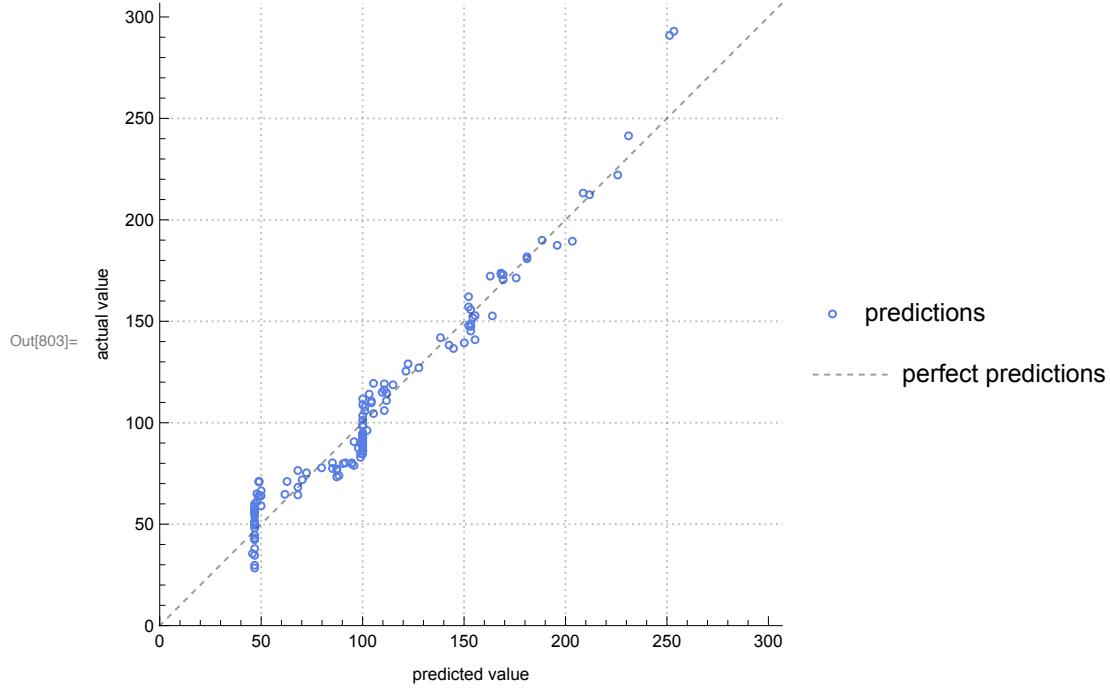
```
In[802]:= Dataset[AssociationMap[
```

```
pm4p[#, ComputeUncertainty → True] &, {"StandardDeviation", "RSquared"}]]
```

Out[802]=

StandardDeviation	$10.5 \pm 0.8$
RSquared	$0.956 \pm 0.010$

```
In[803]:= PredictorMeasurements[pred4p, testingSet2, "ComparisonPlot"]
```



```
In[804]:= PredictorMeasurements[pred4p, testingSet2, "BestPredictedExamples"]
```

```
Out[804]= { {16, 162.11, 2.40113, 0.140013} → 68.1987, {50, 326.436, 5.50825, 1.02686} → 180.832,
{56, 399.629, 4.00636, 1.11738} → 212.396, {24, 432.983, 3.29147, 0.690096} → 127.074,
{46, 360.837, 4.40847, 0.818306} → 181.766,
{26, 213.664, 3.23827, 0.889769} → 104.524,
{22, 305.164, 3.05661, 0.650049} → 110.901, {26, 188.23, 2.63173, 0.506435} → 99.0568,
{51, 290.447, 3.43609, 1.05953} → 170.506, {24, 216.192, 2.76533, 0.178155} → 101.274 }
```

```
In[805]:= PredictorMeasurements[pred4p, testingSet2, "WorstPredictedExamples"]

Out[805]= {{83, 516.857, 4.37087, 1.49007} → 290.869, {77, 610.565, 4.84004, 1.20624} → 292.969,
{20, 121.183, 2.15394, 0.577071} → 71.072, {20, 121.183, 2.12776, 0.174728} → 71.072,
{17, 141.601, 2.16121, 0.107677} → 70.7954, {9, 46.069, 1.18816, 0.388046} → 28.4149,
{16, 142.118, 2.07814, 0.563632} → 65.007, {8, 60.052, 1.27159, 0.223196} → 29.7665,
{20, 176.124, 2.52135, 0.694036} → 78.931, {20, 116.16, 2.17011, 0.738543} → 66.466}
```

## Decision Tree

```
In[806]:= pred5p = Predict[trainingSet2, Method → "DecisionTree"]
```

```
Out[806]= PredictorFunction[  Input type: Mixed (number: 4)
Method: DecisionTree]
```

```
In[807]:= pm5p = PredictorMeasurements[pred5p, testingSet2]
```

```
Out[807]= PredictorMeasurementsObject[  Predictor: DecisionTree
Number of test examples: 135]
```

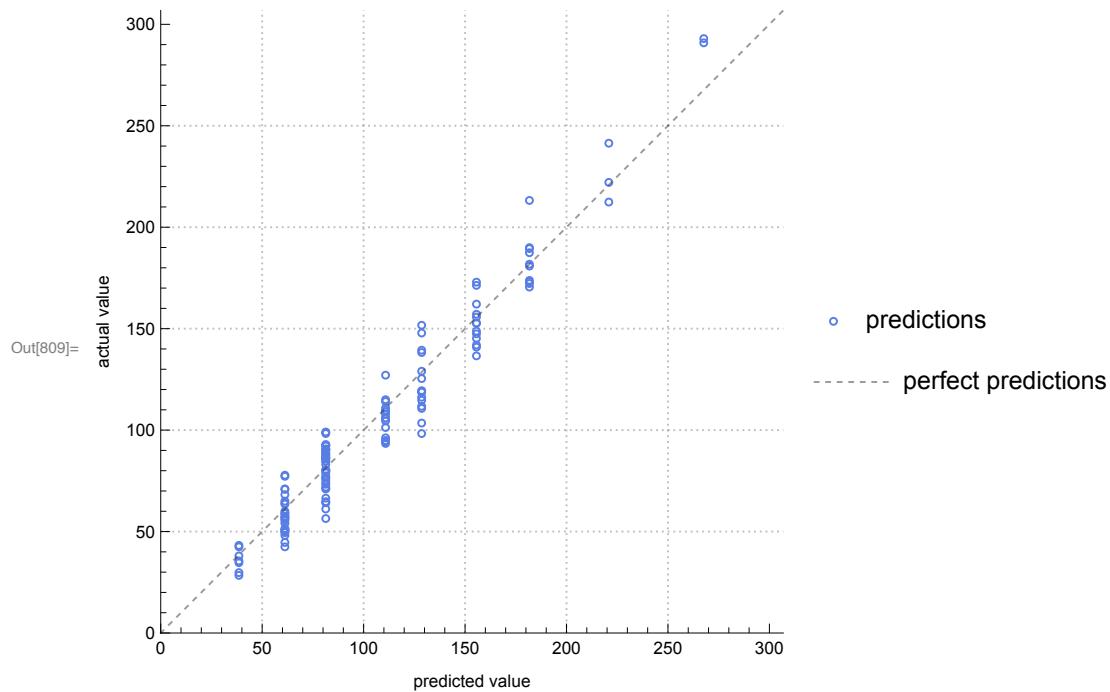
```
In[808]:= Dataset[AssociationMap[
pm5p[#, ComputeUncertainty → True] &, {"StandardDeviation", "RSquared"}]]
```

```
Out[808]= 

|                   |                   |
|-------------------|-------------------|
| StandardDeviation | $11.5 \pm 0.7$    |
| RSquared          | $0.947 \pm 0.011$ |


```

In[809]:= PredictorMeasurements[pred5p, testingSet2, "ComparisonPlot"]



In[810]:= PredictorMeasurements[pred5p, testingSet2, "BestPredictedExamples"]

```
Out[810]= { {23, 135.21, 2.39165, 0.627388} → 80.2468, {23, 135.21, 2.41631, 0.548598} → 80.2468,
{23, 135.21, 2.35493, 0.508427} → 80.2468, {15, 130.099, 2.24907, 0.416276} → 60.0241,
{21, 150.177, 2.45114, 0.249324} → 79.8808,
{24, 163.173, 2.3256, 0.583238} → 82.9543, {21, 150.177, 2.34568, 0.711208} → 79.6184,
{13, 184.035, 1.94439, 0.710411} → 59.0341,
{20, 176.124, 2.52135, 0.694036} → 78.931, {14, 123.115, 2.07701, 0.140488} → 58.878 }
```

In[811]:= PredictorMeasurements[pred5p, testingSet2, "WorstPredictedExamples"]

```
Out[811]= { {54, 442.424, 3.71221, 1.24029} → 213.198, {18, 84.162, 1.59879, 0.569429} → 56.4178,
{31, 156.269, 2.96814, 0.790133} → 98.3144, {19, 100.165, 1.89749, 0.55762} → 61.1168,
{32, 168.284, 2.46708, 0.852928} → 103.479, {13, 76.095, 1.76504, 0.588484} → 42.4707,
{42, 270.376, 3.35943, 0.931283} → 151.681, {77, 610.565, 4.84004, 1.20624} → 292.969,
{83, 516.857, 4.37087, 1.49007} → 290.869, {26, 188.23, 2.63173, 0.506435} → 99.0568 }
```

## Gaussian Process

In[812]:= pred6p = Predict[trainingSet2, Method -> "GaussianProcess"]

```
Out[812]= PredictorFunction[ +  Input type: Mixed (number: 4)
Method: GaussianProcess ]
```

```
In[813]:= pm6p = PredictorMeasurements[pred6p, testingSet2]
```

```
Out[813]= PredictorMeasurementsObject[ + Predictor: GaussianProcess  
Number of test examples: 135 ]
```

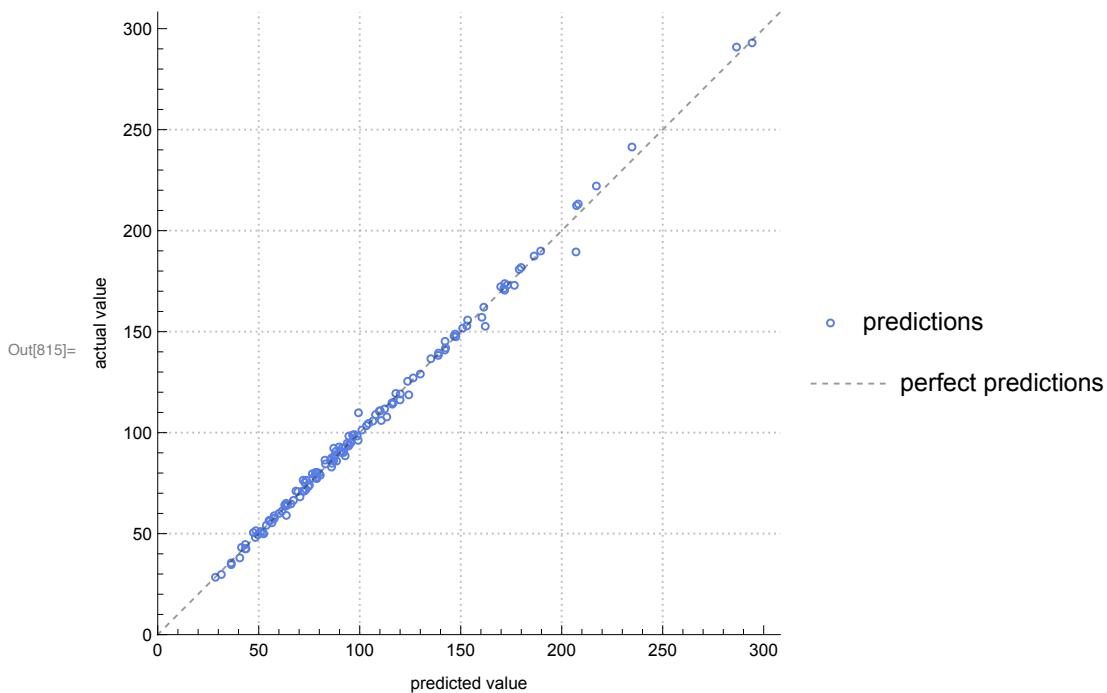
```
In[814]:= Dataset[AssociationMap[
```

```
pm6p[#, ComputeUncertainty → True] &, {"StandardDeviation", "RSquared"}]]
```

```
Out[814]= StandardDeviation
```

StandardDeviation	$2.9 \pm 0.4$
RSquared	$0.9966 \pm 0.0012$

```
In[815]:= PredictorMeasurements[pred6p, testingSet2, "ComparisonPlot"]
```



```
In[816]:= PredictorMeasurements[pred6p, testingSet2, "BestPredictedExamples"]
```

```
Out[816]= { {17, 102.133, 1.90175, 0.68918} → 57.2912, {15, 130.099, 2.24907, 0.416276} → 60.0241,  
{13, 102.089, 1.85542, 0.422599} → 49.4678,  
{14, 88.106, 1.77267, 0.458694} → 48.1163, {21, 150.177, 2.45114, 0.249324} → 79.8808,  
{15, 139.11, 2.29374, 0.0000119294} → 63.9901,  
{35, 313.288, 3.30954, 0.994666} → 147.437, {41, 284.359, 3.34284, 1.07844} → 152.799,  
{15, 109.128, 1.86124, 0.0000194784} → 57.6031,  
{17, 72.151, 1.92643, 0.614753} → 51.0587 }
```

```
In[817]:= PredictorMeasurements[pred6p, testingSet2, "WorstPredictedExamples"]

Out[817]= {{38, 363.223, 4.10357, 0.871585} → 152.656,
           {18, 290.832, 2.74484, 0.762268} → 109.808,
           {30, 221.209, 2.85493, 0.623876} → 107.769,
           {33, 238.305, 2.84447, 0.584026} → 118.703,
           {25, 159.232, 2.46414, 0.827593} → 92.2798,
           {56, 399.629, 4.00636, 1.11738} → 212.396, {51, 416.343, 6.95058, 0.734875} → 189.457,
           {67, 429.604, 4.06698, 1.42492} → 241.398, {26, 240.306, 3.31979, 0.870437} → 106.009,
           {13, 184.035, 1.94439, 0.710411} → 59.0341}
```

## Predicting Specific Surface Area's

```
In[819]:= testingSet2[[1]]

Out[819]= {16, 142.118, 2.07814, 0.563632} → 65.007

In[820]:= fullTestingSet2[[1]]

Out[820]= {C12C(NC(=O)N1)NC(=O)N2, 16, 142.118, 2.07814, 0.563632} → 65.007
```

```
In[821]:= Keys[fullTestingSet2[[1]]][[1]]

Out[821]= C12C(NC(=O)N1)NC(=O)N2

In[822]:= Molecule[Keys[fullTestingSet2[[1]]][[1]]]
```

Out[822]= Molecule[ Formula: C<sub>4</sub>H<sub>6</sub>N<sub>4</sub>O<sub>2</sub>  
Atoms: 10(16) Bonds: 11(17)]

```
In[824]:= pred2p[Keys@testingSet2[[1]]]

Out[824]= 65.0013
```

```
In[832]:= testingSet2[[120]]

Out[832]= {56, 399.629, 4.00636, 1.11738} → 212.396

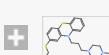
In[833]:= fullTestingSet2[[120]]

Out[833]= {CCSC1=CC2=C(C=C1)SC3=CC=CC=C3N2CCCN4CCN(CC4)C,
           56, 399.629, 4.00636, 1.11738} → 212.396
```

```
In[834]:= Keys[fullTestingSet2[[120]]][[1]]

Out[834]= CCSC1=CC2=C(C=C1)SC3=CC=CC=C3N2CCCN4CCN(CC4)C
```

```
In[835]:= Molecule[%]

Out[835]= Molecule[ Formula: C22H29N3S2  
Atoms: 27(56) Bonds: 30(59)]
```

```
In[837]:= pred2p[Keys@testingSet2[[120]]]
Out[837]= 208.459

In[838]:= testingSet2[[67]]
Out[838]= {23, 135.21, 2.35493, 0.508427} → 80.2468

In[840]:= pred2o[Keys@testingSet2[[67]]]
Out[840]= 80.0597

In[845]:= fullTestingSet2[[67]]
Out[845]= {CCCC1=CC=CC=C1N, 23, 135.21, 2.35493, 0.508427} → 80.2468

In[846]:= Keys[fullTestingSet2[[67]][[1]]]
Out[846]= CCCC1=CC=CC=C1N

In[847]:= Molecule[%]
Out[847]= Molecule[   Formula: C9H13N
Atoms: 10(23) Bonds: 10(23) ]
```

```
In[1221]:= fullTestingSet2[[128]]
Out[1221]= {CN(C)CCCC1(C2=C(CO1)C=C(C=C2)C≡N)C3=CC=C(C=C3)F,
45, 324.399, 3.93421, 1.17359} → 171.338

In[1222]:= pred2p[Keys@testingSet2[[128]]]
Out[1222]= 171.59
```