Interpretability:

Feature extraction:

* For each smiles, a weight matrix is extracted from the attention part of the readout layer from the model. It contains weights for the interconnection of the atoms.
* The attention part consists of a fully connected layer, and a softmax layer.
* The fully connected layer performs weighted sum of the atoms.
* The Softmax function is applied to the weights to keep the total sum to 1.
* For each SMILES, the fragment with maximum weights are extracted. The maximum length of the fragment can be input.
* For each of fragments, the atoms in it have to be above a threshold of 0.5. This is to only include important substructures.

def maxSum(arr, n, k):  
  
 # k must be greater  
 if (n < k):  
 k=n  
 #print(arr)  
  
 # Compute sum of first  
 # window of size k  
 res = 0  
 start=0  
 end=k  
 for i in range(k):  
 if (arr[i] > 0.5):  
 res += arr[i]  
 else:  
 if(end<n):  
 start=start+1  
 end=end+1  
 else:  
 if(k>1):  
 k=k-1  
 else:  
 start=0  
 end=0  
 return start, end  
  
 # Compute sums of remaining windows by  
 # removing first element of previous  
 # window and adding last element of  
 # current window.  
 curr\_sum = res  
 for i in range(k, n):  
 if(arr[i]>0.1):  
 curr\_sum += arr[i] - arr[i - k]  
 if(curr\_sum<res):  
 res = curr\_sum  
 start=i-k+1  
 end=i+1  
  
  
  
  
  
  
 return list(range(start,end))

def processUnit(iMol,start,i,batch\_size,count,tracker,adj\_len,full\_size):  
  
 size = (300, 300)  
 tmp = rdkit.Chem.rdmolfiles.MolFragmentToSmiles(iMol, atomsToUse=start)  
 tmp1=tmp  
 j=0  
 full\_size=full\_size  
  
 start=start  
 while(rdkit.Chem.rdmolfiles.MolFromSmiles(tmp)==None and len(tmp)!=0):  
 j=j+1  
 if(full\_size>=3):  
 full\_size=full\_size-j  
 start = maxSum(tracker,adj\_len,full\_size)  
 else:  
 fig = Draw.MolToFile(iMol, "./amesfirstmodImg4/" + str(i \* batch\_size + count) + '.png', size=size,  
 highlightAtoms=start)  
 return max(tmp1.split('.'), key=len)  
 if(len(start)>0):  
 tmp = rdkit.Chem.rdmolfiles.MolFragmentToSmiles(iMol, atomsToUse=start)  
 #print(tmp)  
 else:  
 fig = Draw.MolToFile(iMol, "./amesfirstmodImg4/" + str(i \* batch\_size + count) + '.png', size=size,  
 highlightAtoms=start)  
 return max(tmp1.split('.'), key=len)  
 fig = Draw.MolToFile(iMol, "./amesfirstmodImg4/" + str(i \* batch\_size + count) + '.png', size=size,  
 highlightAtoms=start)  
 return max(tmp.split('.'), key=len)

SAs in chemical classes:

To allow researchers to have the chemical characterisation of our test dataset. The testdata is classified using Classyfire. Since there are many chemical classes, to focus on the ones concerned by the researchers, a reduction of the chemical classes is performed with an external software(i.e. notepad++) Since a lot of chemical classes are under Benezoid or Bezene and substituted derivatives, subclasses where the researchers have shown interest are taken to other subclasses.

The distribution of the chemical classes is listed in table.

|  |  |
| --- | --- |
| Chemical class | The percentage in the dataset |
| Alcohols and polyols | 0.0079 |
| Benzene and substituted derivatives | 0.1727 |
| Benzoic acids | 0.0327 |
| hydrocarbons | 0.0058 |
| Biphenyls and derivatives | 0.0173 |
| carboxylic acids and derivatives | 0.0815 |
| Organic thiophosphoric acids | 0.0008 |

|  |  |
| --- | --- |
| Chemical class | The percentage in the dataset |
| ethers | 0.0490 |
| Lipids and lipid-like molecules | 0.0393 |
| Organohalogen compounds | 0.0198 |
| Organoheterocyclic | 0.2319 |
| Organonitrogen compounds | 0.0385 |
| Phenylpropanoids and polyketides | 0.0317 |
| phenols | 0.0106 |

Table. The percentages of chemical classes in the test dataset. The other chemical classes not listed in the table include Others:0.0979,Other Benzenoids :0.1060, Other Organic acids and derivatives:0.0243,Other Organic oxygen compounds:0.0320.

The structural alerts are passed in as input and validated by toxtree by chemical classes as listed

|  |  |  |  |
| --- | --- | --- | --- |
| Statistics\Chemical Classes | Alcohols and polyols | Benzene and substituted derivatives | Benzoic acids and derivatives |
| Accuracy | 0.92 | 0.69 | 0.66 |
| MCC | 0.54 | 0.27 | 0.17 |
| Specificity | 0.97 | 0.75 | 0.75 |
| Precision | 0.67 | 0.45 | 0.39 |
| Recall | 0.5 | 0.43 | 0.54 |
| F1 | 0.67 | 0.39 | 0.45 |

|  |  |  |  |
| --- | --- | --- | --- |
| Statistics\Chemical Classes | Carboxylic acids and derivatives | Ethers | Hydrocarbons |
| Accuracy | 0.86 | 0.79 | nah |
| MCC | 0.52 | 0.37 | nah |
| Specificity | 0.93 | 0.92 | nah |
| Precision | 0.65 | 0.61 | nah |
| Recall | 0.56 | 0.41 | nah |
| F1 | 0.65 | 0.61 | nah |

|  |  |  |  |
| --- | --- | --- | --- |
| Statistics\Chemical Classes | Lipids and lipid-like molecules | Organohalogen compounds | Organoheterocyclic |
| Accuracy | 0.77 | 0.78 | 0.63 |
| MCC | 0.18 | 0.57 | 0.16 |
| Specificity | 0.94 | 0.90 | 0.68 |
| Precision | 0.47 | 0.84 | 0.37 |
| Recall | 0.19 | 0.64 | 0.49 |
| F1 | 0.47 | 0.84 | 0.37 |

|  |  |  |  |
| --- | --- | --- | --- |
| Statistics\Chemical Classes | Organonitrogen compounds | phenols | Phenylpropanoids and polyketides |
| Accuracy | 0.82 | 0.68 | 0.71 |
| MCC | 0.58 | 0.38 | 0.20 |
| Specificity | 0.88 | 0.64 | 0.83 |
| Precision | 0.73 | 0.5 | 0.42 |
| Recall | 0.69 | 0.77 | 0.37 |
| F1 | 0.73 | 0.5 | 0.42 |

List the rest as others

Pick 3 classes find struct alerts in graphs and discuss overall model performance

Expert knowledge:

The extracted substructures are inputted in Toxtree. Toxtree is a QSAR model with expert knowledge as its database. This is to check if each substructure consists of at least a subset that can be identified as known structural alerts. The occurrences and the types of the structural alerts in each input are documented and counted.

Visualization:

To visualize the substructure extracted, the model automatically outputs the image of the test data with the substructure identified. In figure, it shows the visualization of different smiles with different substructures. The extraction works on any size of the smiles given as shown.

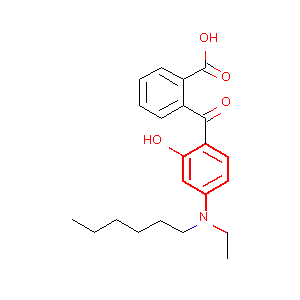
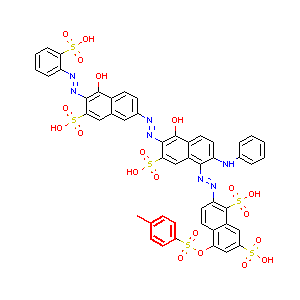
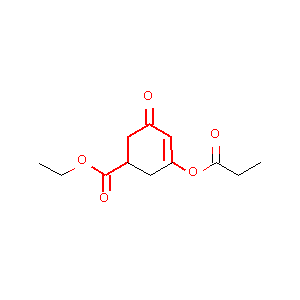
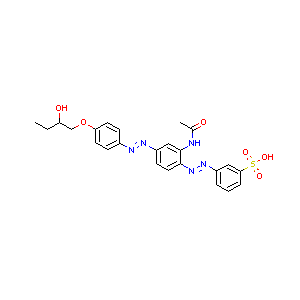


Fig. The image of the structural alerts identified and generated by the model. The smiles listed are O=C(Nc3cc(N=Nc1ccc(OCC(O)CC)cc1)ccc3(N=Nc2cccc(c2)S(=O)(=O)O))C, O=S(=O)(O)c8cc(OS(=O)(=O)c1ccc(cc1)C)c9ccc(N=Nc5c(ccc6c(O)c(N=Nc3ccc4c(O)c(N=Nc2ccccc2S(=O)(=O)O)c(cc4(c3))S(=O)(=O)O)c(cc56)S(=O)(=O)O)Nc7ccccc7)c(c9(c8))S(=O)(=O)O,O=C(O)c2ccccc2(C(=O)c1ccc(cc1(O))N(CC)CCCCCC)

In the figure, it shows the structural alerts identified in Toxtree for smiles O=C5OC2(c4ccc(cc4(Oc1cc(c(cc12)Nc3ccccc3)C))N(CCCC)CCCC)c6ccccc56. The structural alerts identified is QSA18\_Ames.Polycyclic Aromatic Hydrocarbons. In our model nearly the same substructure is identified in the graph for fragment size of 20. For fragment size of 10, the substructure extracted is smaller but the same structural alert can be identified. It can be seen in the graph QSA18\_Ames.Polycyclic Aromatic Hydrocarbons is identified.

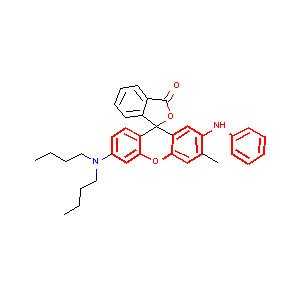
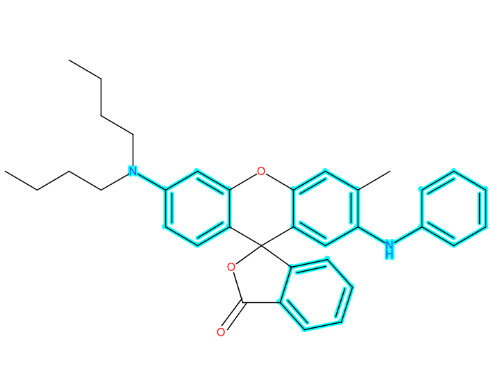
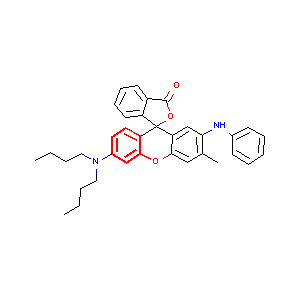


Fig. The image generated by Toxtree is on the top left corner with blue highlight. The other two are substructure identified by our model with fragment size of 10 and fragment size of 20 respectively. The smiles used is O=C5OC2(c4ccc(cc4(Oc1cc(c(cc12)Nc3ccccc3)C))N(CCCC)CCCC)c6ccccc56

In table, there is a part of the list of the structural alert extracted by the model and their corresponding toxtree results. The known structural alerts identified can be shown. The full list can be found in github

|  |  |  |
| --- | --- | --- |
| Smiles extracted by the model | Matched structural alerts in Toxtree | Model Predicted Value |
| Cc1ccc([N+](=O)[O-])cc1 | SA27\_Ames | Mutagenic |
| S/C(Cl)=C(\Cl)CCl | SA8\_Ames | Mutagenic |
| ccccOCCO | SA24\_Ames | Mutagenic |
| c1ccc2ncccc2c1 | SA69\_Ames | Mutagenic |
| Cc1ccc([N+](=O)[O-])cc1 | SA27\_Ames | Mutagenic |
| O=[N+]([O-])c1cccs1 | SA27\_Ames | Mutagenic |
| O=CCO | SA11\_Ames | Mutagenic |
| C\C=N/N | SA13\_Ames | Mutagenic |
| Nc1ccccc1 | SA28\_Ames | Mutagenic |
| COC(=O)[C@H](C)CN=[N+]=[N-] | SA22\_Ames | Mutagenic |
| O=C1OCC(C(Br)Br)=C1Cl | SA65\_Ames | Mutagenic |
| CC=CC(=O)Cl | SA1\_Ames | Mutagenic |
| C/C(=C(/Cl)C(=O)O)C(Cl)Cl | SA8\_Ames | Mutagenic |
| Cc1ccccc1[N+](=O)[O-] | SA27\_Ames | Mutagenic |
| ClCc1ccccc1Cl | SA8\_Ames | Mutagenic |
| COC(=O)[C@H](C)CN=[N+]=[N-] | SA22\_Ames | Mutagenic |
| O=C1OCC(C(Br)Br)=C1Cl | SA8\_Ames, SA65\_Ames | Mutagenic |
| CC=CC(=O)Cl | SA1\_Ames, SA10\_Ames | Mutagenic |
| C/C(=C(/Cl)C(=O)O)C(Cl)Cl | SA8\_Ames | Mutagenic |
| Cc1ccccc1[N+](=O)[O-] | SA27\_Ames | Mutagenic |
| ClCc1ccccc1Cl | SA8\_Ames | Mutagenic |
| CCCl | SA8\_Ames | Mutagenic |
| CCCCC=O | SA11\_Ames | Mutagenic |
| Oc1ccccc1 | None | Nonmutagenic |
| FC(F)(F)c1ccccc1 | None | Nonmutagenic |
| C1=CCCCC1 | None | Nonmutagenic |
| NCc1ccccn1 | None | Nonmutagenic |
| FC(F)c1ccccc1 | None | Nonmutagenic |
| C[N+](C)(C)CCCN | None | Nonmutagenic |
| COC=O | None | Nonmutagenic |
| CCC=C1CCCCC1C | None | Nonmutagenic |
| Oc1ccccc1 | None | Nonmutagenic |
| Clc1ccccc1 | None | Nonmutagenic |
| CCNCCOC(=O)CC | None | Nonmutagenic |
| CCOCCC(=O)C(F)F | None | Nonmutagenic |
| CN=Nc1ccccc1 | None | Nonmutagenic |
| N#Cc1ccc([nH])c(n)c1 | None | Nonmutagenic |
| C=C/c1ccccc1 | None | Nonmutagenic |
| C=C\C1=C(C)CCCC1C | None | Nonmutagenic |
| FCOC(F)(F)C(F)(F)Br | None | Nonmutagenic |
| cccccs | None | Nonmutagenic |
| Cc1ccccc1 | None | Nonmutagenic |
| C | None | Nonmutagenic |
| CSc1ccc(C)cc1 | None | Nonmutagenic |
| CCCc1ccccc1 | None | Nonmutagenic |
| CCCC | None | Nonmutagenic |
| CNCc1ccncc1 | None | Nonmutagenic |
| CNC | None | Nonmutagenic |
| Cc1ccc(O)cc1C | None | Nonmutagenic |
| NCC(F)(F)F | None | Nonmutagenic |
| C/C=C/C | None | Nonmutagenic |

The substructures extracted by the model from the 4800 ames testset are validated with Toxtree. The substructures that are predicted as mutagenic correspond to the known structural alerts in Toxtree. The number of occurrences in the positive substructures for each known structural alert can be found in the table.

There are 22 known structural alerts that can not be found. They are SA12\_Ames, SA15\_Ames, SA18\_Ames, SA19\_Ames, SA26\_Ames, SA28ter\_Ames, SA29\_Ames, SA30\_Ames, SA37\_Ames, SA39\_Ames, SA3\_Ames, SA39\_Ames, SA3\_Ames, SA57\_Ames, SA58\_Ames, SA59\_Ames, SA62\_Ames, SA63\_Ames, SA66\_Ames, SA67\_Ames, SA68\_Ames, SA6\_Ames. It is possible that the testset does not include SMILES with these structural alerts. Upon further research,

Another possible reason is there are many smiles with experimental values that are non mutagenic but identified as mutagenic in Toxtree. They are not counted in the occurrences. This is also the reason why we decide to build a model without apriori knowledge since structural alerts only helps the experts explain why a chemical is toxic, but other mechanisms at organ or cellular level can transform them into other chemicals before they act.

|  |  |
| --- | --- |
| Toxtree Structural Alerts | The number of occurrences identified in the substructure extracted |
| SA10\_Ames | 26 |
| SA11\_Ames | 29 |
| SA12\_Ames | 0 |
| SA13\_Ames | 10 |
| SA14\_Ames | 6 |
| SA15\_Ames | 0 |
| SA16\_Ames | 2 |
| SA18\_Ames | 0 |
| SA19\_Ames | 0 |
| SA1\_Ames | 12 |
| SA21\_Ames | 34 |
| SA22\_Ames | 11 |
| SA23\_Ames | 5 |
| SA24\_Ames | 21 |
| SA25\_Ames | 11 |
| SA26\_Ames | 0 |
| SA27\_Ames | 115 |
| SA28\_Ames | 77 |
| SA28bis\_Ames | 11 |
| SA28ter\_Ames | 0 |
| SA29\_Ames | 0 |
| SA2\_Ames | 8 |
| SA30\_Ames | 0 |
| SA37\_Ames | 0 |
| SA38\_Ames | 8 |
| SA39\_Ames | 0 |
| SA3\_Ames | 0 |
| SA4\_Ames | 6 |
| SA57\_Ames | 0 |
| SA58\_Ames | 0 |
| SA59\_Ames | 0 |
| SA5\_Ames  SA60\_Ames | 2  0 |
| SA61\_Ames | 1 |
| SA62\_Ames | 0 |
| SA63\_Ames | 0 |
| SA64\_Ames | 2 |
| SA65\_Ames | 4 |
| SA66\_Ames | 0 |
| SA67\_Ames | 0 |
| SA68\_Ames | 0 |
| SA69\_Ames | 4 |
| SA6\_Ames | 0 |
| SA7\_Ames | 34 |
| SA8\_Ames | 91 |
| SA9\_Ames | 1 |

Known structural alerts with highest occurrences are SA27\_Ames, SA8\_Ames. The statistics can be seen in table.

|  |  |
| --- | --- |
| Statistics\Structural alerts | SA27\_Ames |
| Accuracy | 0.71 |
| MCC | 0.25 |
| Specificity | 0.09 |
| Precision | 0.70 |
| Recall | 1.0 |
| F1 | 0.70 |

The analysis of whether a structural alert is identified in Toxtree is listed in the table. Since there are many smiles in the Ames test set that have experimental values different from Toxtree, this can be a reason why the accuracy is not high. This is also the reason why we decide to build a model without expert knowledge since structural alerts selected by experts are not enough to determine the toxicity of the chemical.

However, the distribution of the specificity is similar to the model’s overall predictive performance. The model has high specificity, which is 0.81. It is the highest of comparing to other performance parameters. For structural alerts, the highest is also specificity at 0.74, which is understandable since the training data consists of mostly non mutagenic smiles.

|  |  |
| --- | --- |
| Statistics\Structural alerts | All the known structural alerts in Toxtree |
| Accuracy | 0.68 |
| MCC | 0.23 |
| Specificity | 0.74 |
| Precision | 0.38 |
| Recall | 0.51 |
| F1 | 0.38 |