VEGA QSAR MODELS

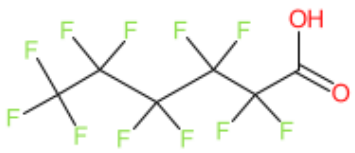
Report

Prediction and Applicability Domain analysis for models:

- Mutagenicity (Ames test) CONSENSUS model 1.0.4
- Mutagenicity (Ames test) model (CAESAR) 2.1.14
- Mutagenicity (Ames test) model (ISS) 1.0.3
- Mutagenicity (Ames test) model (SarPy-IRFMN) 1.0.8
- Mutagenicity (Ames test) model (KNN-Read-Across) 1.0.1
- Mutagenicity (Ames test) model for aromatic amines (CONCERT/IRFMN) 1.0.0
- Carcinogenicity model (CAESAR) 2.1.10
- Carcinogenicity model (ISS) 1.0.3
- Carcinogenicity model (IRFMN-ISSCAN-CGX) 1.0.2
- Carcinogenicity model (IRFMN-Antares) 1.0.2
- Carcinogenicity oral classification model (IRFMN) 1.0.1
- Carcinogenicity oral Slope Factor model (IRFMN) 1.0.1
- Carcinogenicity inhalation classification model (IRFMN) 1.0.1
- Carcinogenicity inhalation Slope Factor model (IRFMN) 1.0.1
- Carcinogenicity in male rat (CORAL) 1.0.0
- Carcinogenicity in female Rat (CORAL) 1.0.0
- Acute Toxicity (LD50) model (KNN) 1.0.0
- BCF model (CAESAR) 2.1.15
- BCF model (Meylan) 1.0.4
- BCF model (Arnot-Gobas) 1.0.1
- BCF model (KNN-Read-Across) 1.1.1

Core version: 1.3.18

Prediction for compound Molecule 0 -



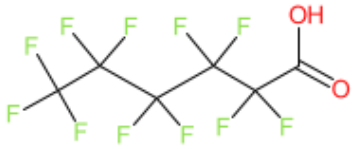
Prediction:

Prediction is NON-Mutagenic with a consensus score of 0.575, based on 4 models.

Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Used models: 4
Predicted Consensus Mutagen activity: NON-Mutagenic
Mutagenic Score: 0
Non-Mutagenic Score: 0.575
Model Caesar assessment: NON-Mutagenic (GOOD reliability)
Model ISS assessment: NON-Mutagenic (LOW reliability)
Model SarPy assessment: NON-Mutagenic (MODERATE reliability)
Model KNN assessment: NON-Mutagenic (MODERATE reliability)
Remarks:
none

1. Prediction Summary

Prediction for compound Molecule 0 -

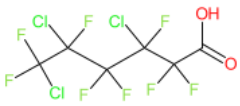


Prediction: Reliability:

Prediction is NON-Mutagenic, the result appears reliable. Anyhow, you should check it through the evaluation of the information given in the following sections.

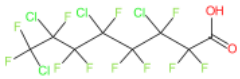
Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
Structural Alerts: -
Reliability: The predicted compound is into the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 2106-54-9
Dataset id:384 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)Cl)Cl)Cl
Similarity: 0.933
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

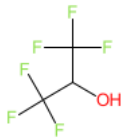


Compound #2

CAS: 2923-68-4
Dataset id:2832 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)Cl)Cl)Cl)Cl
Similarity: 0.871
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

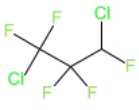
Compound #3

CAS: 335-76-2
Dataset id:3947 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.846
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #4

CAS: 920-66-1
Dataset id:3517 (Training Set)
SMILES: FC(F)(F)C(O)C(F)(F)F
Similarity: 0.719
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #5

CAS: 507-55-1
Dataset id:34 (Training Set)
SMILES: FC(C(F)(F)C(F)(F)Cl)Cl
Similarity: 0.664
Experimental value : NON-Mutagenic
Predicted value : Suspect Mutagenic

Alerts (not found also in the target): SA8 Aliphatic halogens

Compound #6

CAS: 422-56-0
Dataset id:854 (Training Set)
SMILES: FC(F)(F)C(F)(F)C(Cl)Cl
Similarity: 0.659
Experimental value : NON-Mutagenic
Predicted value : Suspect Mutagenic

Alerts (not found also in the target): SA8 Aliphatic halogens

3.2 Applicability Domain:
Measured Applicability Domain Scores

✓	Global AD Index AD index = 0.936 Explanation: The predicted compound is into the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.877 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
✓	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✓	Concordance for similar molecules Concordance index = 1 Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

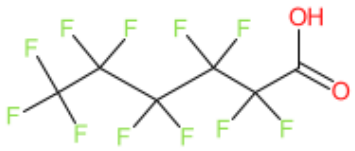
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

1. Prediction Summary

Prediction for compound Molecule 0 -

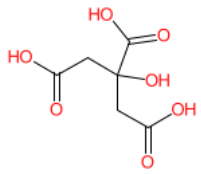


Prediction: Reliability:

Prediction is NON-Mutagenic, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments and 1 infrequent_fragments found)

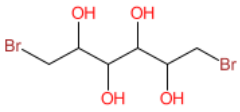
Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
Structural Alerts: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

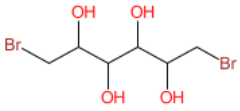
CAS: 77-92-9
Dataset id:829 (Training Set)
SMILES: O=C(O)CC(O)(C(=O)O)CC(=O)O
Similarity: 0.634
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #2

CAS: 10318-26-0
Dataset id:445 (Training Set)
SMILES: OC(CBr)C(O)C(O)C(O)CBr
Similarity: 0.607
Experimental value : Mutagenic
Predicted value : Mutagenic

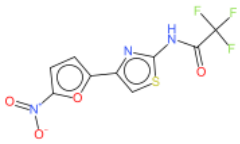
Alerts (not found also in the target): SA8 Aliphatic halogens



Compound #3

CAS: 488-41-5
Dataset id:484 (Training Set)
SMILES: OC(CBr)C(O)C(O)C(O)CBr
Similarity: 0.607
Experimental value : Mutagenic
Predicted value : Mutagenic

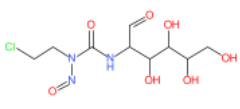
Alerts (not found also in the target): SA8 Aliphatic halogens



Compound #4

CAS: 42011-48-3
Dataset id:763 (Training Set)
SMILES: O=C(Nc1nc(cs1)c2oc(cc2)[N+](=O)[O-])C(F)(F)F
Similarity: 0.606
Experimental value : Mutagenic
Predicted value : Mutagenic

Alerts (not found also in the target): SA27 Nitro aromatic

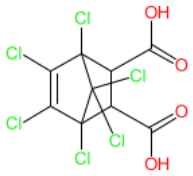


Compound #5

CAS: 54749-90-5
Dataset id:826 (Training Set)
SMILES: O=NN(C(=O)NC(C=O)C(O)C(O)C(O)CO)CCCl
Similarity: 0.604
Experimental value : Mutagenic
Predicted value : Mutagenic

Alerts (not found also in the target): SA8 Aliphatic halogens; SA11 Simple aldehyde; SA21 Alkyl and aryl N-nitroso groups

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #6

CAS: 115-28-6
Dataset id:160 (Training Set)
SMILES: O=C(O)C1C(C(=O)O)C2(C(=C(C1(C2(Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.599
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.341 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.62 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✔	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
⚠	Concordance for similar molecules Concordance index = 0.521 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✖	Atom Centered Fragments similarity check ACF index = 0.51 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments and 1 infrequent_fragments found)..

Symbols explanation:

✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

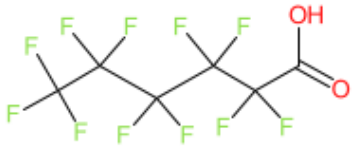
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

	Fragment defined by the SMILES: CC(F)(F)F The fragment has less than 3 occurrences in the model's training set
	Fragment defined by the SMILES: CC(C)(F)F The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

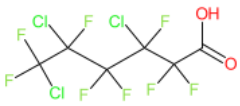


Prediction: Reliability:

Prediction is NON-Mutagenic, but the result shows some critical aspects, which require to be checked:
- Accuracy of prediction for similar molecules found in the training set is not adequate
The following relevant fragments have been found: SM150; SM177

Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
No. alerts for mutagenicity: 0
No. alerts for non-mutagenicity: 2
Structural Alerts: SM150; SM177
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

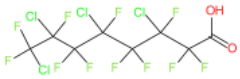


Compound #1

CAS: 2106-54-9
Dataset id:384 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)Cl)Cl)Cl
Similarity: 0.933
Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM150; SM177

Alerts (not found also in the target): SM106



Compound #2

CAS: 2923-68-4
Dataset id:2832 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)Cl)Cl)Cl)Cl
Similarity: 0.871
Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM150; SM177

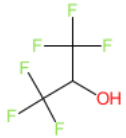
Alerts (not found also in the target): SM106; SM163

Compound #3

CAS: 335-76-2
Dataset id:3947 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.846
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM150; SM177

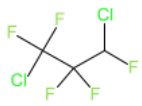
Alerts (not found also in the target): SM143; SM163



Compound #4

CAS: 920-66-1
Dataset id:3517 (Training Set)
SMILES: FC(F)(F)C(O)C(F)(F)F
Similarity: 0.719
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM150



Compound #5

CAS: 507-55-1
Dataset id:34 (Training Set)
SMILES: FC(C(F)(F)C(F)(F)Cl)Cl
Similarity: 0.664
Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM150

Alerts (not found also in the target): SM106

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

Compound #6

CAS: 422-56-0
Dataset id:854 (Training Set)
SMILES: FC(F)(F)C(F)(F)C(Cl)Cl
Similarity: 0.659
Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM150

Alerts (not found also in the target): SM106

3.2 Applicability Domain:
Measured Applicability Domain Scores

⚠	Global AD Index AD index = 0.703 Explanation: The predicted compound could be out of the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.877 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
✗	Accuracy of prediction for similar molecules Accuracy index = 0.317 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✓	Concordance for similar molecules Concordance index = 1 Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

✓ The feature has a good assessment, model is reliable regarding this aspect.

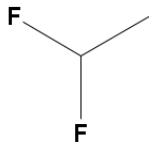
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

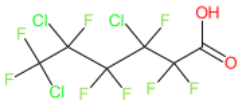
(Molecule 0) Reasoning on fragments/structural alerts - 1 of 2:.

Fragment found: SM150



Sarpy alert n. 150 for NON-Mutagenicity, defined by SMARTS: C(F)(F)C

Following, the most similar compounds from the model's dataset having the same fragment.

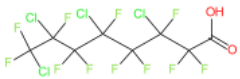


CAS: 2106-54-9
Dataset id:384 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)Cl)Cl)Cl
Similarity: 0.933

Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM150; SM177

Alerts (not found also in the target): SM106



CAS: 2923-68-4
Dataset id:2832 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)Cl)Cl)Cl)Cl
Similarity: 0.871

Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM150; SM177

Alerts (not found also in the target): SM106; SM163

CAS: 335-76-2
Dataset id:3947 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.846

Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

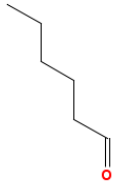
Alerts (found also in the target): SM150; SM177

Alerts (not found also in the target): SM143; SM163

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

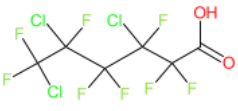
(Molecule 0) Reasoning on fragments/structural alerts - 2 of 2:.

Fragment found: SM177



Sarpy alert n. 177 for NON-Mutagenicity, defined by SMARTS: C(=O)CCCCC

Following, the most similar compounds from the model's dataset having the same fragment.

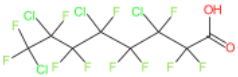


CAS: 2106-54-9
Dataset id:384 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)Cl)Cl)Cl
Similarity: 0.933

Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM150; SM177

Alerts (not found also in the target): SM106



CAS: 2923-68-4
Dataset id:2832 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)Cl)Cl)Cl)Cl
Similarity: 0.871

Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM150; SM177

Alerts (not found also in the target): SM106; SM163

CAS: 335-76-2
Dataset id:3947 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.846

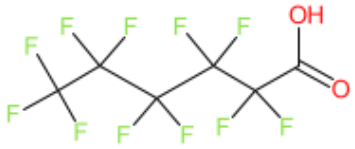
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM150; SM177

Alerts (not found also in the target): SM143; SM163

1. Prediction Summary

Prediction for compound Molecule 0 -

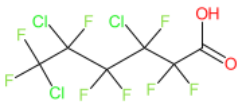


Prediction: Reliability:

Prediction is NON-Mutagenic, but the result shows some critical aspects, which require to be checked:
- Accuracy of prediction for similar molecules found in the training set is not optimal
- some similar molecules found in the training set have experimental values that disagree with the predicted value

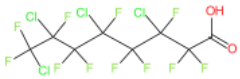
Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
Molecules used for prediction: 4
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 2106-54-9
Dataset id:2157 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)Cl)Cl)Cl
Similarity: 0.933
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

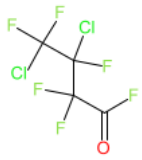


Compound #2

CAS: 2923-68-4
Dataset id:2605 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)C(F)(C(F)(F)Cl)Cl)Cl)Cl
Similarity: 0.871
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

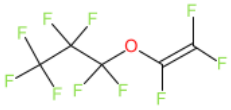
Compound #3

CAS: 335-76-2
Dataset id:2785 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.846
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #4

CAS: 678-06-8
Dataset id:5701 (Training Set)
SMILES: O=C(F)C(F)(F)C(F)(C(F)(F)Cl)Cl
Similarity: 0.779
Experimental value : Mutagenic
Predicted value : NON-Mutagenic



Compound #5

CAS: 1623-05-8
Dataset id:5707 (Training Set)
SMILES: FC(F)=C(F)OC(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.777
Experimental value : Mutagenic
Predicted value : NON-Mutagenic

Compound #6

CAS: 423-39-2
Dataset id:3062 (Training Set)
SMILES: FC(F)(F)C(F)(F)C(F)(F)C(F)(F)I
Similarity: 0.762
Experimental value : NON-Mutagenic
Predicted value : Mutagenic

3.2 Applicability Domain:
Measured Applicability Domain Scores

⚠	Global AD Index AD index = 0.807 Explanation: The predicted compound could be out of the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.837 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.778 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 0.778 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

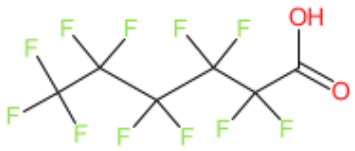
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

1. Prediction Summary

Prediction for compound Molecule 0 -

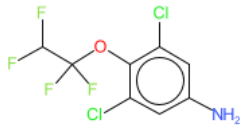


Prediction: Reliability:

Prediction is NA, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not adequate
- similar molecules found in the training set have experimental values that disagree with the predicted value
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)

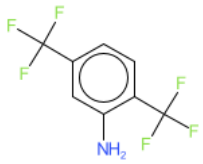
Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Mutagen activity: NA
Structural Alerts: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



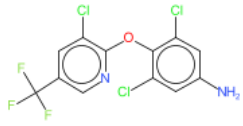
Compound #1

CAS: N.A.
Dataset id:5582 (Training Set)
SMILES: FC(F)C(F)(F)Oc1c(cc(N)cc1Cl)Cl
Similarity: 0.613
Experimental value : NON-Mutagenic
Predicted value : Mutagenic



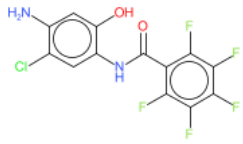
Compound #2

CAS: N.A.
Dataset id:7541 (Training Set)
SMILES: FC(F)(F)c1ccc(c(N)c1)C(F)(F)F
Similarity: 0.611
Experimental value : NON-Mutagenic
Predicted value : Mutagenic



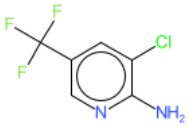
Compound #3

CAS: N.A.
Dataset id:4773 (Training Set)
SMILES: FC(F)(F)c2cnc(Oc1c(cc(N)cc1Cl)Cl)c(c2)Cl
Similarity: 0.561
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



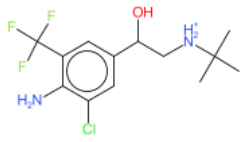
Compound #4

CAS: N.A.
Dataset id:4500 (Training Set)
SMILES: O=C(Nc1cc(c(N)cc1(O))Cl)c2c(F)c(F)c(F)c(F)c2(F)
Similarity: 0.557
Experimental value : NON-Mutagenic
Predicted value : Mutagenic



Compound #5

CAS: N.A.
Dataset id:4740 (Training Set)
SMILES: FC(F)(F)c1cnc(N)c(c1)Cl
Similarity: 0.557
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #6

CAS: N.A.
Dataset id:5067 (Training Set)
SMILES: FC(F)(F)c1cc(cc(c1(N))Cl)C(O)C[NH2+]C(C)(C)C
Similarity: 0.555
Experimental value : NON-Mutagenic
Predicted value : Mutagenic

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.612 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✖	Accuracy of prediction for similar molecules Accuracy index = 0 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✖	Concordance for similar molecules Concordance index = 0 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✖	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..

Symbols explanation:

✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

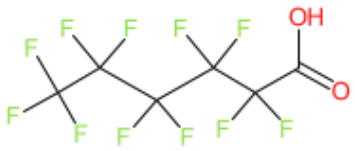
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

Fragment defined by the SMILES: CC(C)(F)F
The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

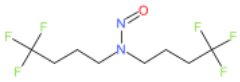


Prediction: Reliability:

Prediction is Carcinogen, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)
- predicted substance falls into a neuron that is populated by no compounds of the training set

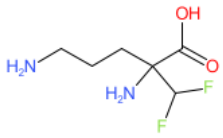
Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Carcinogen activity: Carcinogen
P(Carcinogen): 0.736
P(NON-Carcinogen): 0.264
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 83335-32-4
Dataset id:541 (Training Set)
SMILES: O=NN(CCCC(F)(F)F)CCCC(F)(F)F
Similarity: 0.687
Experimental value : Carcinogen
Predicted value : Carcinogen

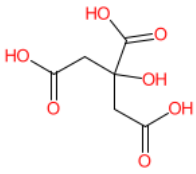


Compound #2

CAS: 70052-12-9
Dataset id:245 (Training Set)
SMILES: O=C(O)C(N)(CCCN)C(F)F
Similarity: 0.676
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen

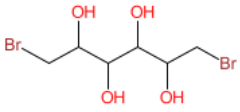
Compound #3

CAS: 625-89-8
Dataset id:576 (Test Set)
SMILES: O=NN(CC(F)(F)F)CC(F)(F)F
Similarity: 0.658
Experimental value : NON-Carcinogen
Predicted value : Carcinogen



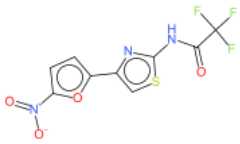
Compound #4

CAS: 77-92-9
Dataset id:173 (Training Set)
SMILES: O=C(O)CC(O)(C(=O)O)CC(=O)O
Similarity: 0.634
Experimental value : NON-Carcinogen
Predicted value : Carcinogen



Compound #5

CAS: 10318-26-0
Dataset id:213 (Training Set)
SMILES: OC(CBr)C(O)C(O)C(O)CBr
Similarity: 0.607
Experimental value : Carcinogen
Predicted value : NON-Carcinogen



Compound #6

CAS: 42011-48-3
Dataset id:774 (Training Set)
SMILES: O=C(Nc1nc(cs1)c2oc(cc2)[N+](=O)[O-])C(F)(F)F
Similarity: 0.606
Experimental value : Carcinogen
Predicted value : Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.209 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.681 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✓	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
⚠	Concordance for similar molecules Concordance index = 0.507 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✗	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..
✓	Model class assignment reliability Pos/Non-Pos difference = 0.473 Explanation: model class assignment is well defined..
✗	Neural map neurons concordance Neurons concordance = 0.5 Explanation: predicted substance falls into a neuron that is populated by no compounds of the training set..

Symbols explanation:

✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

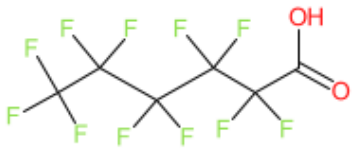
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

Fragment defined by the SMILES: CC(C)(F)F
The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

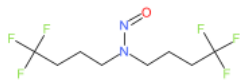


Prediction: Reliability:

Prediction is Carcinogen, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not optimal
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)
The following alerts have been found: SA43 Perfluorooctanoic acid (PFOA)

Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Carcinogen activity: Carcinogen
Structural Alerts: SA43 Perfluorooctanoic acid (PFOA)
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

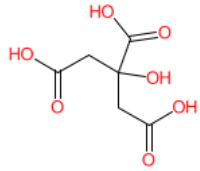
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 83335-32-4
Dataset id:496 (Training Set)
SMILES: O=NN(CCCC(F)(F)F)CCCC(F)(F)F
Similarity: 0.687
Experimental value : Carcinogen
Predicted value : Carcinogen

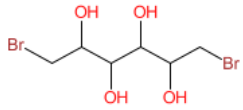
Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups



Compound #2

CAS: 77-92-9
Dataset id:829 (Training Set)
SMILES: O=C(O)CC(O)(C(=O)O)CC(=O)O
Similarity: 0.634
Experimental value : NON-Carcinogen
Predicted value : Carcinogen

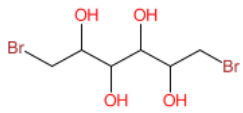
Alerts (not found also in the target): SA41 Substituted n-alkylcarboxylic acids



Compound #3

CAS: 10318-26-0
Dataset id:445 (Training Set)
SMILES: OC(CBr)C(O)C(O)C(O)CBr
Similarity: 0.607
Experimental value : Carcinogen
Predicted value : Carcinogen

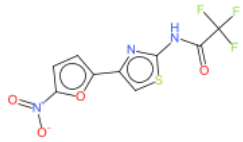
Alerts (not found also in the target): SA8 Aliphatic halogens



Compound #4

CAS: 488-41-5
Dataset id:484 (Training Set)
SMILES: OC(CBr)C(O)C(O)C(O)CBr
Similarity: 0.607
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): SA8 Aliphatic halogens

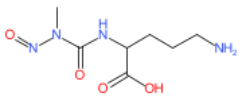


Compound #5

CAS: 42011-48-3
Dataset id:763 (Training Set)
SMILES: O=C(Nc1nc(cs1)c2oc(cc2)[N+](=O)[O-])C(F)(F)F
Similarity: 0.606
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): SA27 Nitro aromatic

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #6

CAS: 63642-17-1
Dataset id:757 (Training Set)
SMILES: O=NN(C(=O)NC(C(=O)O)CCCN)C
Similarity: 0.604
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.356 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.659 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.534 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 0.534 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✖	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..

Symbols explanation:

✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on fragments/structural alerts :.

Fragment found: SA43 Perfluorooctanoic acid (PFOA)

Perfluorooctanoic acid (PFOA)

No compounds with the same fragment have been found int the model's dataset.

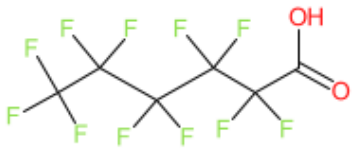
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

Fragment defined by the SMILES: CC(C)(F)F
The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

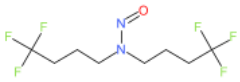


Prediction: Reliability:

Prediction is Possible NON-Carcinogen, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not optimal
- similar molecules found in the training set have experimental values that disagree with the predicted value
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)

Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Carcinogenic activity: Possible NON-Carcinogen
No. alerts for carcinogenicity: 0
Structural Alerts: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

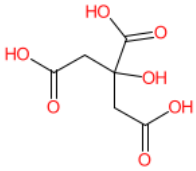
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

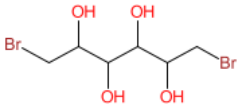
CAS: 83335-32-4
Dataset id:409 (Training Set)
SMILES: O=NN(CCCC(F)(F)F)CCCC(F)(F)F
Similarity: 0.687
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 1; Carcinogenity alert no. 14; Carcinogenity alert no. 27



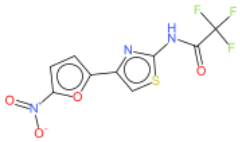
Compound #2

CAS: 77-92-9
Dataset id:745 (Training Set)
SMILES: O=C(O)CC(O)(C(=O)O)CC(=O)O
Similarity: 0.634
Experimental value : NON-Carcinogen
Predicted value : Possible NON-Carcinogen



Compound #3

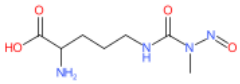
CAS: 10318-26-0
Dataset id:369 (Training Set)
SMILES: OC(CBr)C(O)C(O)C(O)CBr
Similarity: 0.607
Experimental value : Carcinogen
Predicted value : Possible NON-Carcinogen



Compound #4

CAS: 42011-48-3
Dataset id:596 (Training Set)
SMILES: O=C(Nc1nc(cs1)c2oc(cc2)[N+](=O)[O-])C(F)(F)F
Similarity: 0.606
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 2; Carcinogenity alert no. 19

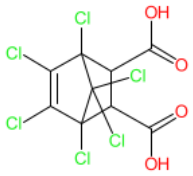


Compound #5

CAS: 63642-17-1
Dataset id:590 (Training Set)
SMILES: O=NN(C(=O)NCCCC(N)C(=O)O)C
Similarity: 0.601
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 3; Carcinogenity alert no. 14; Carcinogenity alert no. 28

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #6

CAS: 115-28-6
Dataset id:127 (Training Set)
SMILES: O=C(O)C1C(C(=O)O)C2(C(=C(C1(C2(Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.599
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 25; Carcinogenity alert no. 39

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.331 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.639 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.7 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✖	Concordance for similar molecules Concordance index = 0.326 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✖	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..

Symbols explanation:

✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

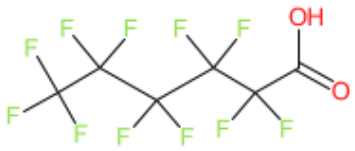
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

Fragment defined by the SMILES: CC(C)(F)F
The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

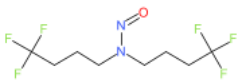


Prediction: Reliability:

Prediction is Carcinogen, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not adequate
- similar molecules found in the training set have experimental values that disagree with the predicted value
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)
The following relevant fragments have been found: Carcinogenity alert no. 125

Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Carcinogenic activity: Carcinogen
No. alerts for carcinogenicity: 1
Structural Alerts: Carcinogenity alert no. 125
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

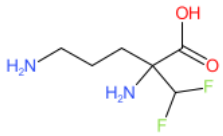


Compound #1

CAS: 83335-32-4
Dataset id:541 (Training Set)
SMILES: O=NN(CCCC(F)(F)F)CCCC(F)(F)F
Similarity: 0.687
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 125

Alerts (not found also in the target): Carcinogenity alert no. 5; Carcinogenity alert no. 8; Carcinogenity alert no. 10; Carcinogenity alert no. 15; Carcinogenity alert no. 50; Carcinogenity alert no. 51; Carcinogenity alert no. 53; Carcinogenity alert no. 54; Carcinogenity alert no. 55; Carcinogenity alert no. 63



Compound #2

CAS: 70052-12-9
Dataset id:245 (Training Set)
SMILES: O=C(O)C(N)(CCCN)C(F)F
Similarity: 0.676
Experimental value : NON-Carcinogen
Predicted value : Carcinogen

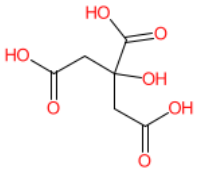
Alerts (found also in the target): Carcinogenity alert no. 125

Compound #3

CAS: 625-89-8
Dataset id:576 (Test Set)
SMILES: O=NN(CC(F)(F)F)CC(F)(F)F
Similarity: 0.658
Experimental value : NON-Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 125

Alerts (not found also in the target): Carcinogenity alert no. 8; Carcinogenity alert no. 50; Carcinogenity alert no. 55; Carcinogenity alert no. 63

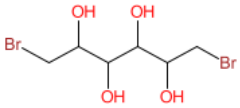


Compound #4

CAS: 77-92-9
Dataset id:173 (Training Set)
SMILES: O=C(O)CC(O)(C(=O)O)CC(=O)O
Similarity: 0.634
Experimental value : NON-Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 21

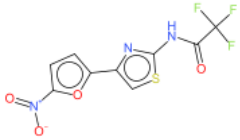
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #5

CAS: 10318-26-0
Dataset id:213 (Training Set)
SMILES: OC(CBr)C(O)C(O)C(O)CBr
Similarity: 0.607
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 58; Carcinogenity alert no. 59



Compound #6

CAS: 42011-48-3
Dataset id:774 (Training Set)
SMILES: O=C(Nc1nc(cs1)c2oc(cc2)[N+](=O)[O-])C(F)(F)F
Similarity: 0.606
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 125

Alerts (not found also in the target): Carcinogenity alert no. 63; Carcinogenity alert no. 64; Carcinogenity alert no. 90; Carcinogenity alert no. 108; Carcinogenity alert no. 117; Carcinogenity alert no. 123

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.289 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.673 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✖	Accuracy of prediction for similar molecules Accuracy index = 0.344 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✖	Concordance for similar molecules Concordance index = 0.344 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✖	Atom Centered Fragments similarity check ACF index = 0.6 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments found)..

Symbols explanation:

✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

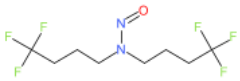
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on fragments/structural alerts :.

Fragment found: Carcinogenity alert no. 125

Structural alert for carcinogenity defined by the SMARTS: CCF

Following, the most similar compounds from the model's dataset having the same fragment.

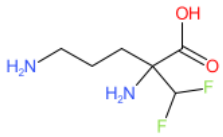


CAS: 83335-32-4
Dataset id:541 (Training Set)
SMILES: O=NN(CCCC(F)(F)F)CCCC(F)(F)F
Similarity: 0.687

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 125

Alerts (not found also in the target): Carcinogenity alert no. 5; Carcinogenity alert no. 8; Carcinogenity alert no. 10; Carcinogenity alert no. 15; Carcinogenity alert no. 50; Carcinogenity alert no. 51; Carcinogenity alert no. 53; Carcinogenity alert no. 54; Carcinogenity alert no. 55; Carcinogenity alert no. 63



CAS: 70052-12-9
Dataset id:245 (Training Set)
SMILES: O=C(O)C(N)(CCCN)C(F)F
Similarity: 0.676

Experimental value : NON-Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 125

CAS: 625-89-8
Dataset id:576 (Test Set)
SMILES: O=NN(CC(F)(F)F)CC(F)(F)F
Similarity: 0.658

Experimental value : NON-Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 125

Alerts (not found also in the target): Carcinogenity alert no. 8; Carcinogenity alert no. 50; Carcinogenity alert no. 55; Carcinogenity alert no. 63

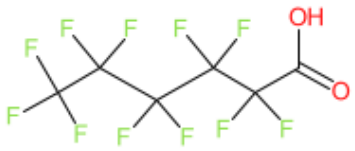
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

Fragment defined by the SMILES: CC(C)(F)F
The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -



Prediction: Reliability:

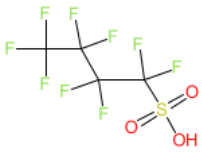
Prediction is NON-Carcinogen, the result appears reliable. Anyhow, you should check it through the evaluation of the information given in the following sections.

Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Oral Carcinogenic class: NON-Carcinogen
Reliability: The predicted compound is into the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

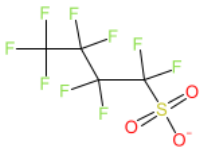
Compound #1

CAS: 335-67-1
Dataset id:629 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.906
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



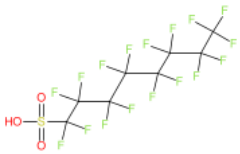
Compound #2

CAS: 375-73-5
Dataset id:627 (Training Set)
SMILES: O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.792
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



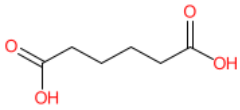
Compound #3

CAS: 29420-49-3
Dataset id:647 (Test Set)
SMILES: O=S(=O)([O-])C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.781
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



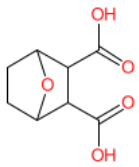
Compound #4

CAS: 1763-23-1
Dataset id:628 (Training Set)
SMILES: O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.753
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #5

CAS: 124-04-9
Dataset id:541 (Training Set)
SMILES: O=C(O)CCCCC(=O)O
Similarity: 0.63
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #6

CAS: 145-73-3
Dataset id:494 (Training Set)
SMILES: O=C(O)C2C1OC(CC1)C2(C(=O)O)
Similarity: 0.601
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✓	Global AD Index AD index = 0.915 Explanation: The predicted compound is into the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.838 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
✓	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✓	Concordance for similar molecules Concordance index = 1 Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

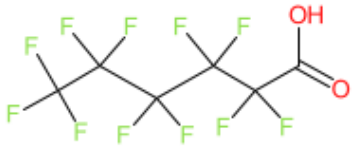
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

1. Prediction Summary

Prediction for compound Molecule 0 -

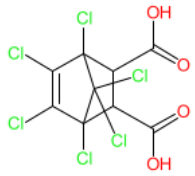


Prediction: Reliability:

Prediction is -1.46, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- No similar compounds with known experimental value in the training set have been found
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (2 unknown fragments and 1 infrequent_fragments found)

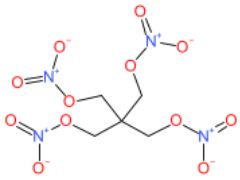
Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Oral Carcinogenicity SF (log form) [log(1/(mg/kg-day))]: -1.46
Predicted Oral Carcinogenicity SF [1/(mg/kg-day)]: 0.0344
Experimental value [1/(mg/kg-day)]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



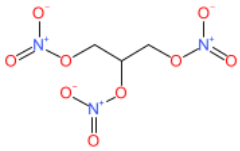
Compound #1

CAS: 115-28-6
Dataset id:64 (Test Set)
SMILES: O=C(O)C1C(C(=O)O)C2(C(=C(C1(C2(Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.599
Experimental value : -1.04
Predicted value : -0.711



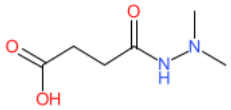
Compound #2

CAS: 78-11-5
Dataset id:253 (Training Set)
SMILES: O=[N+]([O-])OCC(CO[N+](=O)[O-])(CO[N+](=O)[O-])CO[N+](=O)[O-]
Similarity: 0.596
Experimental value : -2.4
Predicted value : -2.004



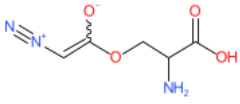
Compound #3

CAS: 55-63-0
Dataset id:218 (Training Set)
SMILES: O=[N+]([O-])OCC(O[N+](=O)[O-])CO[N+](=O)[O-]
Similarity: 0.59
Experimental value : -1.77
Predicted value : -1.69



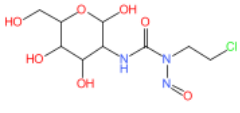
Compound #4

CAS: 1596-84-5
Dataset id:88 (Training Set)
SMILES: O=C(O)CCC(=O)NN(C)C
Similarity: 0.582
Experimental value : -1.74
Predicted value : -2.103



Compound #5

CAS: 115-02-6
Dataset id:27 (Training Set)
SMILES: N#[N+]C=C([O-])OCC(N)C(=O)O
Similarity: 0.579
Experimental value : 1.04
Predicted value : 0.917



Compound #6

CAS: 54749-90-5
Dataset id:78 (Training Set)
SMILES: O=NN(C(=O)NC1C(O)OC(CO)C(O)C1(O))CCCl
Similarity: 0.578
Experimental value : 2.38
Predicted value : 2.294

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.203 Explanation: The predicted compound is outside the Applicability Domain of the model.
✖	Similar molecules with known experimental value Similarity index = 0.597 Explanation: No similar compounds with known experimental value in the training set have been found..
✔	Accuracy of prediction for similar molecules Accuracy index = 0.363 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✔	Concordance for similar molecules Concordance index = 0.68 Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
✔	Maximum error of prediction among similar molecules Max error index = 0.396 Explanation: the maximum error in prediction of similar molecules found in the training set has a low value, considering the experimental variability..
✔	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✖	Atom Centered Fragments similarity check ACF index = 0.34 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (2 unknown fragments and 1 infrequent_fragments found)..

Symbols explanation:

✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

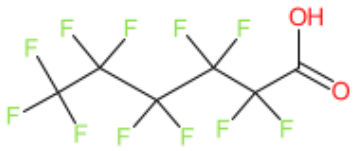
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

	Fragment defined by the SMILES: CF The fragment has less than 3 occurrences in the model's training set
	Fragment defined by the SMILES: CC(C)(F)F The fragment has never been found in the model's training set
	Fragment defined by the SMILES: CC(F)(F)F The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -



Prediction: Reliability:

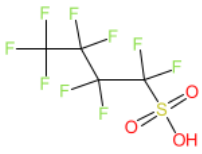
Prediction is Carcinogen, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Accuracy of prediction for similar molecules found in the training set is not adequate
- similar molecules found in the training set have experimental values that disagree with the predicted value
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)

Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Inhalation Carcinogenic class: Carcinogen
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

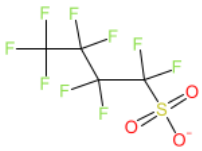
Compound #1

CAS: 335-67-1
Dataset id:615 (Test Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.906
Experimental value : NON-Carcinogen
Predicted value : Carcinogen



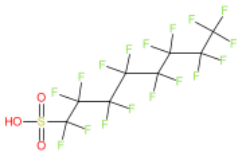
Compound #2

CAS: 375-73-5
Dataset id:613 (Training Set)
SMILES: O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.792
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



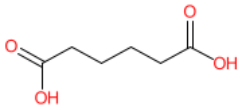
Compound #3

CAS: 29420-49-3
Dataset id:635 (Test Set)
SMILES: O=S(=O)([O-])C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.781
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



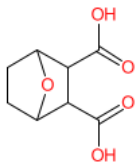
Compound #4

CAS: 1763-23-1
Dataset id:614 (Training Set)
SMILES: O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.753
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #5

CAS: 124-04-9
Dataset id:517 (Training Set)
SMILES: O=C(O)CCCCC(=O)O
Similarity: 0.63
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #6

CAS: 145-73-3
Dataset id:466 (Training Set)
SMILES: O=C(O)C2C1OC(CC1)C2(C(=O)O)
Similarity: 0.601
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0 Explanation: The predicted compound is outside the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.838 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
✗	Accuracy of prediction for similar molecules Accuracy index = 0.46 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✗	Concordance for similar molecules Concordance index = 0 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
⚠	Atom Centered Fragments similarity check ACF index = 0.85 Explanation: some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)..

Symbols explanation:

✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

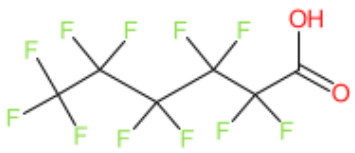
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

Fragment defined by the SMILES: CC(C)(F)F
The fragment has less than 3 occurrences in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

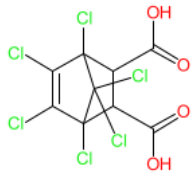


Prediction: Reliability:

Prediction is -0.36, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- No similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not adequate
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (3 unknown fragments found)

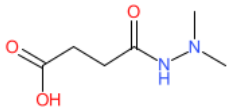
Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted Inhalation Carcinogenicity SF (log form) [log(1/(mg/kg-day))]: -0.36
Predicted Inhalation Carcinogenicity SF [1/(mg/kg-day)]: 0.4403
Experimental value [1/(mg/kg-day)]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



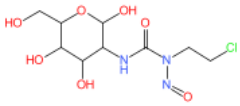
Compound #1

CAS: 115-28-6
Dataset id:53 (Test Set)
SMILES: O=C(O)C1C(C(=O)O)C2(C(=C(C1(C2(Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.599
Experimental value : -1.04
Predicted value : 1.623



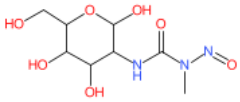
Compound #2

CAS: 1596-84-5
Dataset id:74 (Training Set)
SMILES: O=C(O)CCC(=O)NN(C)C
Similarity: 0.582
Experimental value : -1.75
Predicted value : -0.976



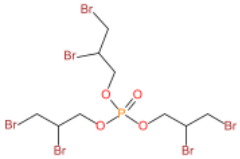
Compound #3

CAS: 54749-90-5
Dataset id:64 (Training Set)
SMILES: O=NN(C(=O)NC1C(O)OC(CO)C(O)C1(O))CCCl
Similarity: 0.578
Experimental value : 2.38
Predicted value : 1.702



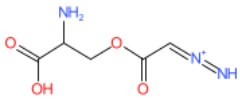
Compound #4

CAS: 18883-66-4
Dataset id:234 (Test Set)
SMILES: O=NN(C(=O)NC1C(O)OC(CO)C(O)C1(O))C
Similarity: 0.578
Experimental value : 2.04
Predicted value : 0.989



Compound #5

CAS: 126-72-7
Dataset id:257 (Training Set)
SMILES: O=P(OCC(CBr)Br)(OCC(CBr)Br)OCC(CBr)Br
Similarity: 0.578
Experimental value : 0.36
Predicted value : 0.569



Compound #6

CAS: 115-02-6
Dataset id:24 (Training Set)
SMILES: O=C(O)C(N)COC(=O)C=[N+]=N
Similarity: 0.577
Experimental value : 1.04
Predicted value : -0.661

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.236 Explanation: The predicted compound is outside the Applicability Domain of the model.
✗	Similar molecules with known experimental value Similarity index = 0.59 Explanation: No similar compounds with known experimental value in the training set have been found..
✗	Accuracy of prediction for similar molecules Accuracy index = 1.719 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
⚠	Concordance for similar molecules Concordance index = 1.039 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✗	Maximum error of prediction among similar molecules Max error index = 2.663 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✗	Atom Centered Fragments similarity check ACF index = 0.4 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (3 unknown fragments found)..

Symbols explanation:

✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

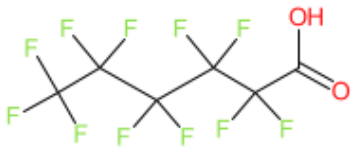
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

	Fragment defined by the SMILES: CC(C)(F)F The fragment has never been found in the model's training set
	Fragment defined by the SMILES: CC(F)(F)F The fragment has never been found in the model's training set
	Fragment defined by the SMILES: CF The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

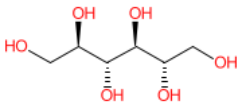


Prediction: Reliability:

Prediction is 2.9086, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- No similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not optimal
- similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (2 unknown fragments and 1 infrequent_fragments found)

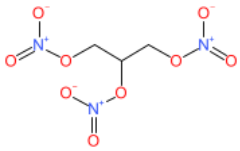
Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted TD50 [-log(mg/kg bw/day)]: 2.9086
Predicted TD50 [mg/kg bw/day]: 0.0012
Experimental TD50 [mg/kg bw/day]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



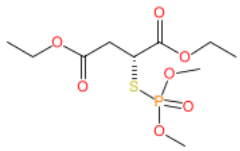
Compound #1

CAS: N.A.
Dataset id:162 (Test Set)
SMILES: C([C@@H]([C@@H]([C@@H]([C@@H](CO)O)O)O)O)O
Similarity: 0.592
Experimental value : -4.255
Predicted value : -3.919



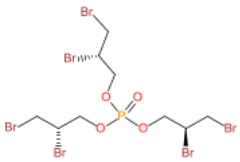
Compound #2

CAS: N.A.
Dataset id:24 (Training Set)
SMILES: C(C(CON(=O)=O)ON(=O)=O)ON(=O)=O
Similarity: 0.59
Experimental value : -2.344
Predicted value : -0.701



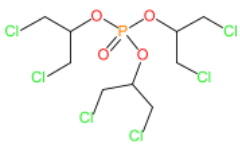
Compound #3

CAS: N.A.
Dataset id:38 (Training Set)
SMILES: CCOC(=O)[C@@H](CC(=O)OCC)SP(=O)(OC)OC
Similarity: 0.583
Experimental value : -3.193
Predicted value : -3.699



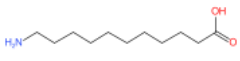
Compound #4

CAS: N.A.
Dataset id:12 (Training Set)
SMILES: P(=O)(OC[C@@H](CBr)Br)(OC[C@@H](CBr)Br)OC[C@@H](CBr)Br
Similarity: 0.578
Experimental value : -1.535
Predicted value : -1.665



Compound #5

CAS: N.A.
Dataset id:18 (Training Set)
SMILES: P(=O)(OC(CCl)CCl)(OC(CCl)CCl)OC(CCl)CCl
Similarity: 0.578
Experimental value : -2.158
Predicted value : -2.535



Compound #6

CAS: N.A.
Dataset id:34 (Training Set)
SMILES: C(CCCCCCCCCC(=O)O)N
Similarity: 0.577
Experimental value : -3.041
Predicted value : -2.357

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.201 Explanation: The predicted compound is outside the Applicability Domain of the model.
✗	Similar molecules with known experimental value Similarity index = 0.591 Explanation: No similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.99 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✗	Concordance for similar molecules Concordance index = 6.208 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✗	Maximum error of prediction among similar molecules Max error index = 1.643 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✗	Atom Centered Fragments similarity check ACF index = 0.34 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (2 unknown fragments and 1 infrequent_fragments found)..

Symbols explanation:

✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

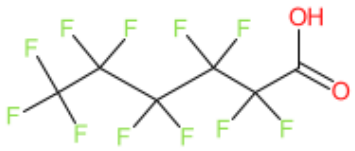
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

	Fragment defined by the SMILES: CF The fragment has less than 3 occurrences in the model's training set
	Fragment defined by the SMILES: CC(C)(F)F The fragment has never been found in the model's training set
	Fragment defined by the SMILES: CC(F)(F)F The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -

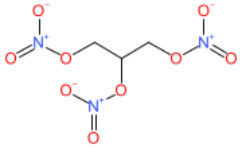


Prediction: Reliability:

Prediction is -4.8952, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- No similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not adequate
- similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (2 unknown fragments and 1 infrequent_fragments found)

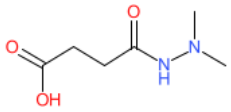
Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted TD50 [-log(mg/kg bw/day)]: -4.8952
Predicted TD50 [mg/kg bw/day]: 78563.51
Experimental TD50 [mg/kg bw/day]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



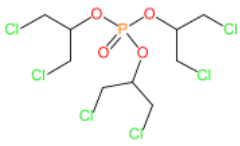
Compound #1

CAS: N.A.
Dataset id:94 (Training Set)
SMILES: C(C(CON(=O)=O)ON(=O)=O)ON(=O)=O
Similarity: 0.59
Experimental value : -2.517
Predicted value : -6.738



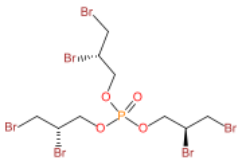
Compound #2

CAS: N.A.
Dataset id:111 (Training Set)
SMILES: C(=O)(O)CCC(=O)NN(C)C
Similarity: 0.582
Experimental value : -3.984
Predicted value : -2.932



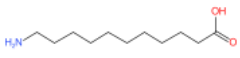
Compound #3

CAS: N.A.
Dataset id:92 (Training Set)
SMILES: P(=O)(OC(CCl)CCl)(OC(CCl)CCl)OC(CCl)CCl
Similarity: 0.578
Experimental value : -2.314
Predicted value : -1.883



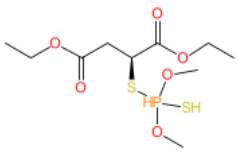
Compound #4

CAS: N.A.
Dataset id:106 (Training Set)
SMILES: P(=O)(OC[C@@H](CBr)Br)(OC[C@@H](CBr)Br)OC[C@@H](CBr)Br
Similarity: 0.578
Experimental value : -3.45
Predicted value : -3.074



Compound #5

CAS: N.A.
Dataset id:117 (Training Set)
SMILES: C(CCCCCCCCCC(=O)O)N
Similarity: 0.577
Experimental value : -4.649
Predicted value : -4.782



Compound #6

CAS: N.A.
Dataset id:107 (Training Set)
SMILES: P(S)(S[C@H](C(=O)OCC)CC(=O)OCC)(OC)OC
Similarity: 0.576
Experimental value : -3.588
Predicted value : -2.078

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.199 Explanation: The predicted compound is outside the Applicability Domain of the model.
✗	Similar molecules with known experimental value Similarity index = 0.586 Explanation: No similar compounds with known experimental value in the training set have been found..
✗	Accuracy of prediction for similar molecules Accuracy index = 2.636 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✗	Concordance for similar molecules Concordance index = 1.645 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✗	Maximum error of prediction among similar molecules Max error index = 4.221 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✗	Atom Centered Fragments similarity check ACF index = 0.34 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (2 unknown fragments and 1 infrequent_fragments found)..

Symbols explanation:

✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

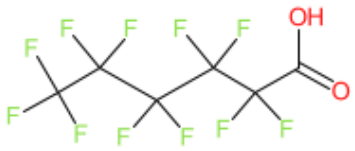
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

	Fragment defined by the SMILES: CF The fragment has less than 3 occurrences in the model's training set
	Fragment defined by the SMILES: CC(C)(F)F The fragment has never been found in the model's training set
	Fragment defined by the SMILES: CC(F)(F)F The fragment has never been found in the model's training set

1. Prediction Summary

Prediction for compound Molecule 0 -



Prediction: Reliability:

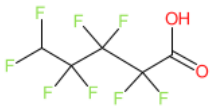
Prediction is 988.01 mg/kg, the result appears reliable. Anyhow, you should check it through the evaluation of the information given in the following sections.

Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted log LD50 [log(mmol/Kg)]: 0.498
Predicted log LD50 [mg/Kg]: 988.01
Molecules used for prediction: 3
Experimental value [mg/Kg]: -
Reliability: The predicted compound is into the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

Compound #1

CAS: N.A.
Dataset id:4123 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)F
Similarity: 0.952
Experimental value : 0.62
Predicted value : 0.397



Compound #2

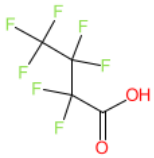
CAS: N.A.
Dataset id:3530 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)F
Similarity: 0.909
Experimental value : 0.85
Predicted value : 0.786

Compound #3

CAS: N.A.
Dataset id:4690 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.906
Experimental value : 0
Predicted value : -0.017

Compound #4

CAS: N.A.
Dataset id:4922 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)F
Similarity: 0.878
Experimental value : 0.27
Predicted value : 0.008



Compound #5

CAS: N.A.
Dataset id:3430 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.862
Experimental value : 0.82
Predicted value : 0.808

Compound #6

CAS: N.A.
Dataset id:418 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.846
Experimental value : -0.96
Predicted value : 0.405

3.2 Applicability Domain:
Measured Applicability Domain Scores

✓	Global AD Index AD index = 1 Explanation: The predicted compound is into the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.92 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
✓	Accuracy of prediction for similar molecules Accuracy index = 0.101 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✓	Concordance for similar molecules Concordance index = 0.324 Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
✓	Maximum error of prediction among similar molecules Max error index = 0.223 Explanation: the maximum error in prediction of similar molecules found in the training set has a low value, considering the experimental variability..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

1. Prediction Summary

Prediction for compound Molecule 0 -

Prediction: Reliability:

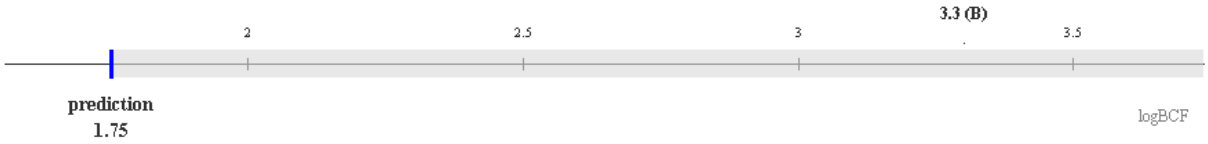
Prediction is 1.75 log(L/kg), but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Only moderately similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not adequate
- similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability
Warning: the prediction may be not fully reliable due to the presence of one or more fragments related to model outliers.
The following relevant fragments have been found: 10 F atoms in the molecule (SO 10); Carbonyl residue (SR 02); COOH group (PG 01)

Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted BCF [log(L/kg)]: 1.75
Predicted BCF [L/kg]: 57
Predicted BCF from sub-model 1 (HM) [log(L/kg)]: 1.72
Predicted BCF from sub-model 2 (GA) [log(L/kg)]: 2.08
Predicted LogP (MLogP): 3.11
Structural Alerts: 10 F atoms in the molecule (SO 10); Carbonyl residue (SR 02); COOH group (PG 01)
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

2. Possible Use and Uncertainty

Threshold 3.3 (bioaccumulative)

Following, a chart showing the predicted value together with its conservative confidence interval for safe classification.
For the threshold logBCF = 3.3, the current compound can not be associated (due to its Applicability Domain index value) to any conservative interval.
No safe classification can be done.



Threshold 3.7 (very bioaccumulative)

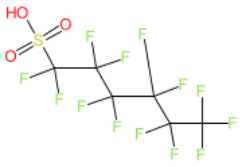
Following, a chart showing the predicted value together with its conservative confidence interval for safe classification.
For the threshold logBCF = 3.7, the current compound can not be associated (due to its Applicability Domain index value) to any conservative interval.
No safe classification can be done.

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

Compound #1

CAS: 335-67-1
Dataset id:56 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.906
Experimental value : 3.12
Predicted value : 2.534

Alerts (found also in the target): 10 F atoms in the molecule (SO 10); Carbonyl residue (SR 02); COOH group (PG 01)

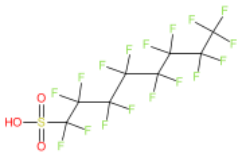


Compound #2

CAS: 355-46-4
Dataset id:55 (Training Set)
SMILES: O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.807
Experimental value : 3.6
Predicted value : 1.585

Alerts (found also in the target): 10 F atoms in the molecule (SO 10)

Alerts (not found also in the target): SO3H group (PG 02)

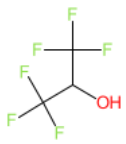


Compound #3

CAS: 1763-23-1
Dataset id:57 (Training Set)
SMILES: O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.753
Experimental value : 3.73
Predicted value : 1.697

Alerts (found also in the target): 10 F atoms in the molecule (SO 10)

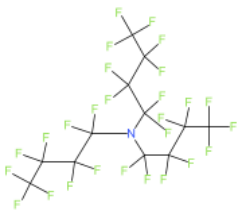
Alerts (not found also in the target): SO3H group (PG 02)



Compound #4

CAS: 920-66-1
Dataset id:263 (Training Set)
SMILES: FC(F)(F)C(O)C(F)(F)F
Similarity: 0.719
Experimental value : 0.3
Predicted value : 0.601

Alerts (not found also in the target): OH group (PG 06)



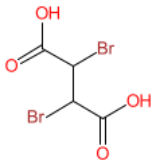
Compound #5

CAS: 311-89-7
Dataset id:419 (Training Set)
SMILES:
FC(F)(F)C(F)(F)C(F)(F)C(F)(F)N(C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.649
Experimental value : 1.42
Predicted value : 0.879

Alerts (found also in the target): 10 F atoms in the molecule (SO 10)

Alerts (not found also in the target): Tertiary amine (SR 05)

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #6

CAS: 526-78-3
Dataset id:71 (Training Set)
SMILES: O=C(O)C(C(C(=O)O)Br)Br
Similarity: 0.618
Experimental value : 0.81
Predicted value : 0.227

Alerts (found also in the target): Carbonyl residue (SR 02); COOH group (PG 01)

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.75 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.848 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✖	Accuracy of prediction for similar molecules Accuracy index = 1.301 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✖	Concordance for similar molecules Concordance index = 1.608 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✖	Maximum error of prediction among similar molecules Max error index = 2.015 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✔	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✔	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on fragments/structural alerts - 1 of 3:.

Fragment found: 10 F atoms in the molecule (SO 10)

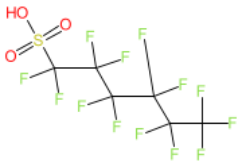
Compounds with ten or more F atoms fall into a chemical category that results out of the applicability domain of the model.

Following, the most similar compounds from the model's dataset having the same fragment.

CAS: 335-67-1
Dataset id:56 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.906

Experimental value : 3.12
Predicted value : 2.534

Alerts (found also in the target): 10 F atoms in the molecule (SO 10); Carbonyl residue (SR 02); COOH group (PG 01)

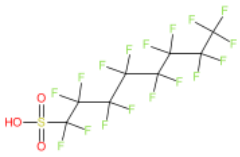


CAS: 355-46-4
Dataset id:55 (Training Set)
SMILES: O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.807

Experimental value : 3.6
Predicted value : 1.585

Alerts (found also in the target): 10 F atoms in the molecule (SO 10)

Alerts (not found also in the target): SO3H group (PG 02)



CAS: 1763-23-1
Dataset id:57 (Training Set)
SMILES: O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.753

Experimental value : 3.73
Predicted value : 1.697

Alerts (found also in the target): 10 F atoms in the molecule (SO 10)

Alerts (not found also in the target): SO3H group (PG 02)

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on fragments/structural alerts - 2 of 3:.

Fragment found: Carbonyl residue (SR 02)

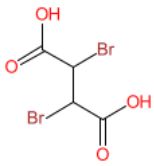
This chemical contains a carbonyl residue. This residue has been found to be present in a very large (112) number of non-bioaccumulative compounds, even when the logP value was higher than 3.

Following, the most similar compounds from the model's dataset having the same fragment.

CAS: 335-67-1
Dataset id:56 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.906

Experimental value : 3.12
Predicted value : 2.534

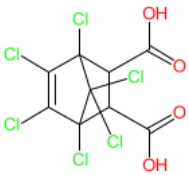
Alerts (found also in the target): 10 F atoms in the molecule (SO 10); Carbonyl residue (SR 02); COOH group (PG 01)



CAS: 526-78-3
Dataset id:71 (Training Set)
SMILES: O=C(O)C(C(C(=O)O)Br)Br
Similarity: 0.618

Experimental value : 0.81
Predicted value : 0.227

Alerts (found also in the target): Carbonyl residue (SR 02); COOH group (PG 01)



CAS: 115-28-6
Dataset id:282 (Training Set)
SMILES: O=C(O)C1C(C(=O)O)C2(C(=C(C1(C2(Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.599

Experimental value : 0.32
Predicted value : 0.261

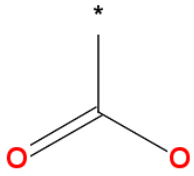
Alerts (found also in the target): Carbonyl residue (SR 02); COOH group (PG 01)

Alerts (not found also in the target): 6 Cl atoms in the molecule (SO 01)

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

(Molecule 0) Reasoning on fragments/structural alerts - 3 of 3:.

Fragment found: COOH group (PG 01)



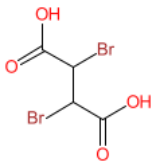
This chemical contains a COOH polar group.The presence of polar groups increases hydrophilicity, related to lower values of BCF.

Following, the most similar compounds from the model's dataset having the same fragment.

CAS: 335-67-1
Dataset id:56 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.906

Experimental value : 3.12
Predicted value : 2.534

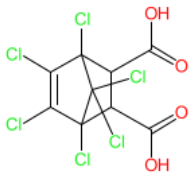
Alerts (found also in the target): 10 F atoms in the molecule (SO 10); Carbonyl residue (SR 02); COOH group (PG 01)



CAS: 526-78-3
Dataset id:71 (Training Set)
SMILES: O=C(O)C(C(C(=O)O)Br)Br
Similarity: 0.618

Experimental value : 0.81
Predicted value : 0.227

Alerts (found also in the target): Carbonyl residue (SR 02); COOH group (PG 01)



CAS: 115-28-6
Dataset id:282 (Training Set)
SMILES: O=C(O)C1C(C(=O)O)C2(C(=C(C1(C2(Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.599

Experimental value : 0.32
Predicted value : 0.261

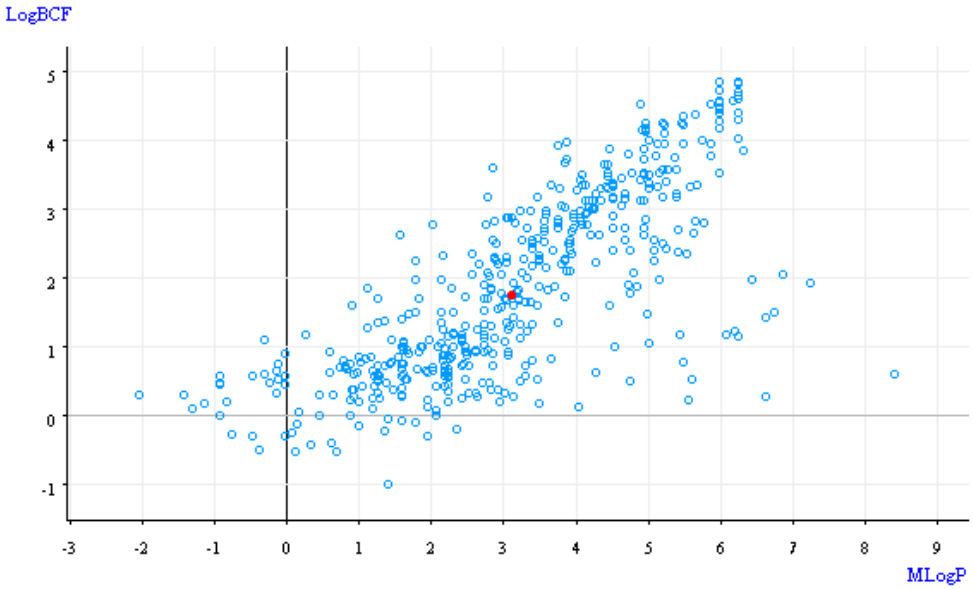
Alerts (found also in the target): Carbonyl residue (SR 02); COOH group (PG 01)

Alerts (not found also in the target): 6 Cl atoms in the molecule (SO 01)

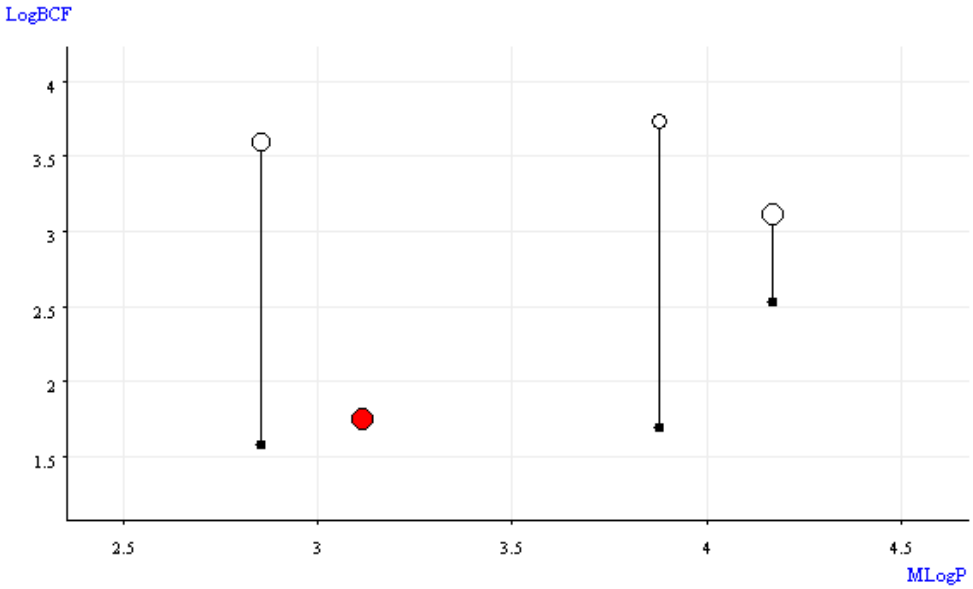
4.2 Reasoning:
Analysis of Molecular Descriptors

Descriptor name: MLogP
Descriptor name: LogP is directly correlated to the logBCF value.

Following, a scatterplot of MLogP against response values; experimental values are reported for the training set, predicted value for the studied compound. Light blue dots represent values of compounds from training set, red dot is the value of the studied compound.

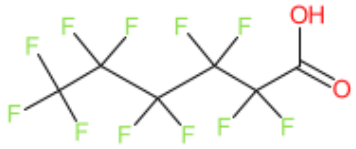


Following, a scatterplot of MLogP against response values only for 3 most similar compounds in the training set. Red dot is the value of the studied compound, black outlined circles represents experimental values of compounds from training set, black dots represents predicted value of the same compound; the size of the circle is proportional to the similarity to the studied compound.



1. Prediction Summary

Prediction for compound Molecule 0 -

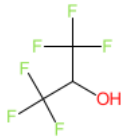


Prediction: Reliability:

Prediction is 0.5 log(L/kg), but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- No similar compounds with known experimental value in the training set have been found
- reliability of logP value used by the model is not adequate
- a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments and 1 infrequent_fragments found)

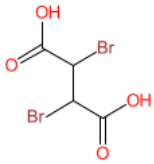
Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted BCF [log(L/kg)]: 0.5
Predicted BCF [L/kg]: 3
Predicted LogP (Meylan/Kowwin): 4.37
Predicted LogP reliability: Low
MW: 313.73
Ionic compound: yes
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



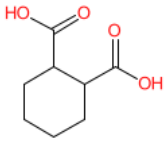
Compound #1

CAS: 920-66-1
Dataset id:117 (Training Set)
SMILES: FC(F)(F)C(O)C(F)(F)F
Similarity: 0.719
Experimental value : 0.4
Predicted value : 0.762



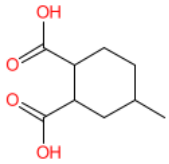
Compound #2

CAS: 526-78-3
Dataset id:27 (Training Set)
SMILES: O=C(O)C(C(C(=O)O)Br)Br
Similarity: 0.618
Experimental value : 0.81
Predicted value : 0.5



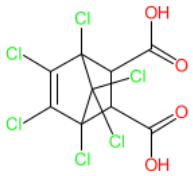
Compound #3

CAS: 1687-30-5
Dataset id:34 (Training Set)
SMILES: O=C(O)C1CCCCC1(C(=O)O)
Similarity: 0.606
Experimental value : 0.3
Predicted value : 0.5



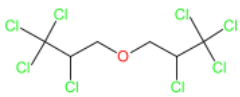
Compound #4

CAS: 57567-84-7
Dataset id:59 (Training Set)
SMILES: O=C(O)C1CCC(C)CC1(C(=O)O)
Similarity: 0.603
Experimental value : 0.38
Predicted value : 0.5



Compound #5

CAS: 115-28-6
Dataset id:17 (Training Set)
SMILES: O=C(O)C1C(C(=O)O)C2(C(=C(C1(C2(Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.599
Experimental value : 0.32
Predicted value : 0.5



Compound #6

CAS: 127-90-2
Dataset id:406 (Training Set)
SMILES: O(CC(C(Cl)(Cl)Cl)Cl)CC(C(Cl)(Cl)Cl)Cl
Similarity: 0.596
Experimental value : 3.28
Predicted value : 3.034

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.337 Explanation: The predicted compound is outside the Applicability Domain of the model.
✗	Similar molecules with known experimental value Similarity index = 0.662 Explanation: No similar compounds with known experimental value in the training set have been found..
✓	Accuracy of prediction for similar molecules Accuracy index = 0.336 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✓	Concordance for similar molecules Concordance index = 0.205 Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
✓	Maximum error of prediction among similar molecules Max error index = 0.362 Explanation: the maximum error in prediction of similar molecules found in the training set has a low value, considering the experimental variability..
✗	Reliability of logP prediction LogP reliability = 0 Explanation: reliability of logP value used by the model is not adequate..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the defined range..
✗	Atom Centered Fragments similarity check ACF index = 0.51 Explanation: a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 unknown fragments and 1 infrequent_fragments found)..

Symbols explanation:

✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

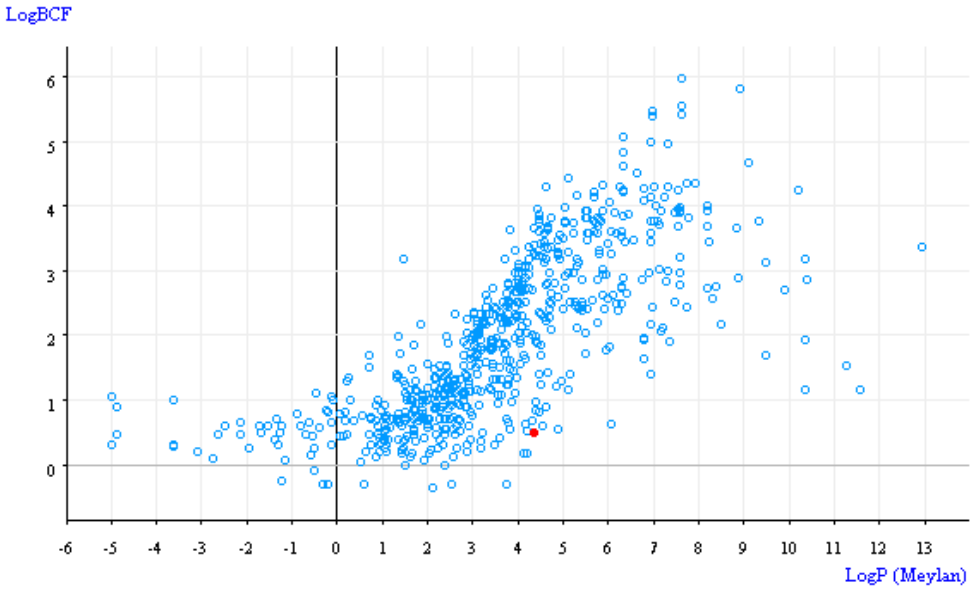
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:

	Fragment defined by the SMILES: CC(F)(F)F The fragment has less than 3 occurrences in the model's training set
	Fragment defined by the SMILES: CC(C)(F)F The fragment has never been found in the model's training set

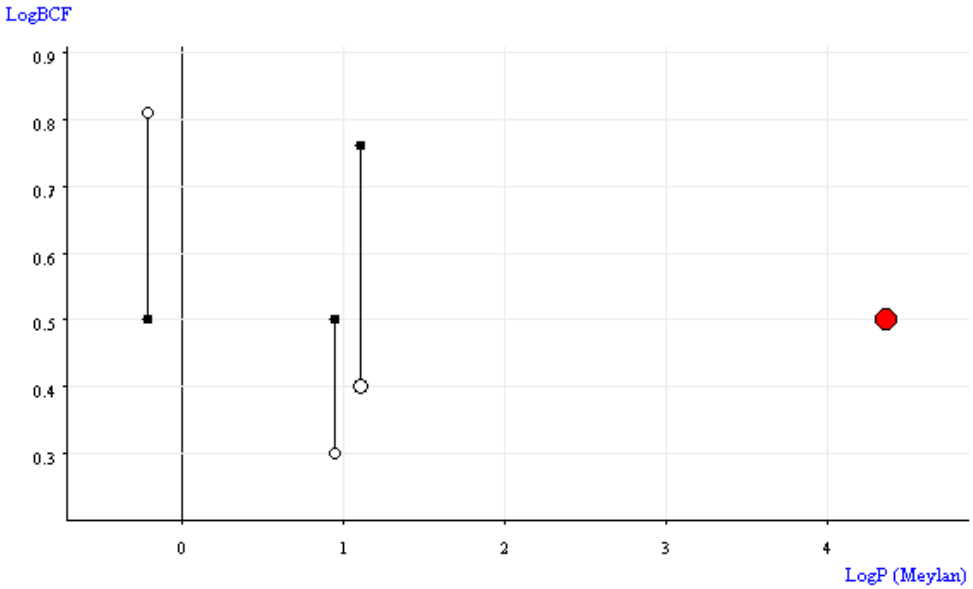
4.2 Reasoning:
Analysis of Molecular Descriptors

Descriptor name: LogP (Meylan)
Descriptor name: LogP is directly correlated to the logBCF value.

Following, a scatterplot of LogP (Meylan) against response values; experimental values are reported for the training set, predicted value for the studied compound. Light blue dots represent values of compounds from training set, red dot is the value of the studied compound.

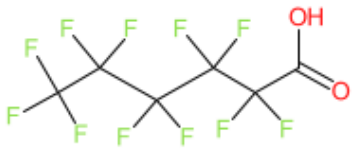


Following, a scatterplot of LogP (Meylan) against response values only for 3 most similar compounds in the training set. Red dot is the value of the studied compound, black outlined circles represents experimental values of compounds from training set, black dots represents predicted value of the same compound; the size of the circle is proportional to the similarity to the studied compound.



1. Prediction Summary

Prediction for compound Molecule 0 -



Prediction: Reliability:

Prediction is 3.16 log(L/kg), but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Accuracy of prediction for similar molecules found in the training set is not adequate
- similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability
- reliability of logP value used by the model is not adequate

Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted BCF (up) [log(L/kg)]: 3.16
Predicted BCF (up) [L/kg]: 1458
Predicted BCF (low) [log(L/kg)]: 3.04
Predicted BCF (low) [L/kg]: 1107
Predicted BCF (mid) [log(L/kg)]: 3.08
Predicted BCF (mid) [L/kg]: 1203
Predicted LogP (Meylan/Kowwin): 4.37
Predicted LogP reliability: Low
Predicted kM (Meylan): 0.97
Predicted kM reliability: Low
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

Compound #1

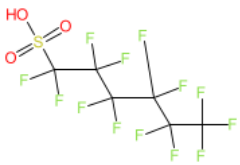
CAS: 335-67-1
Dataset id:644 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.906
Experimental value : 0.977
Predicted value : 4

Compound #2

CAS: 335-76-2
Dataset id:288 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.846
Experimental value : 3.04
Predicted value : 3.236

Compound #3

CAS: 2058-94-8
Dataset id:665 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.821
Experimental value : 3.72
Predicted value : 2.565



Compound #4

CAS: 3871-99-6
Dataset id:430 (Training Set)
SMILES: O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.807
Experimental value : 1.62
Predicted value : 2.989

Compound #5

CAS: 307-55-1
Dataset id:413 (Training Set)
SMILES:
O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.798
Experimental value : 4.373
Predicted value : 1.714

Compound #6

CAS: 376-06-7
Dataset id:163 (Training Set)
SMILES:
O=C(O)C(F)(F)F
Similarity: 0.756
Experimental value : 4.44
Predicted value : 0.194

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.75 Explanation: The predicted compound is outside the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.873 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
✗	Accuracy of prediction for similar molecules Accuracy index = 1.61 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✗	Concordance for similar molecules Concordance index = 1.155 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✗	Maximum error of prediction among similar molecules Max error index = 3.023 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✗	Reliability of logP prediction LogP reliability = 0 Explanation: reliability of logP value used by the model is not adequate..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

✓ The feature has a good assessment, model is reliable regarding this aspect.

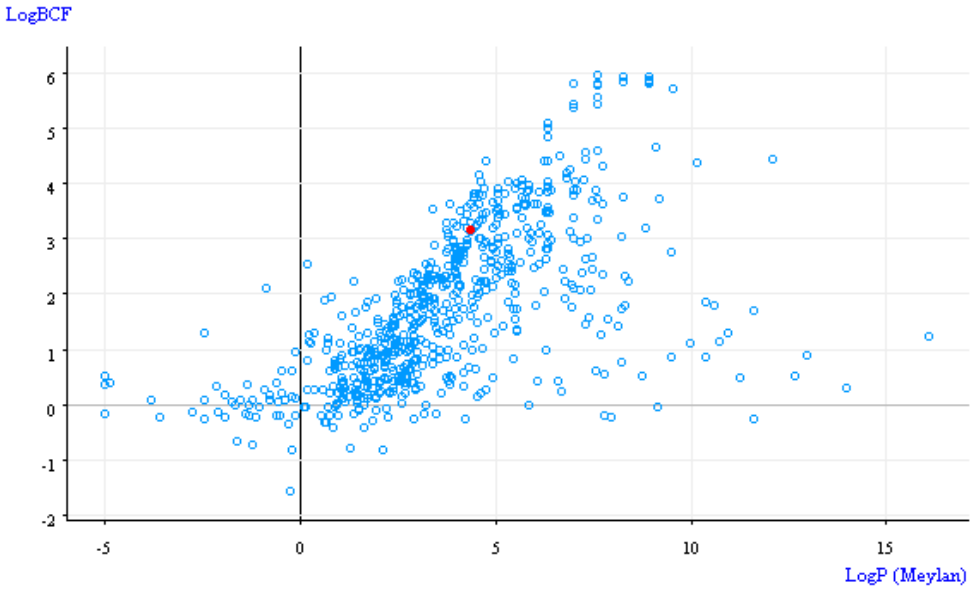
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✗ The feature has a bad assessment, model is not reliable regarding this aspect.

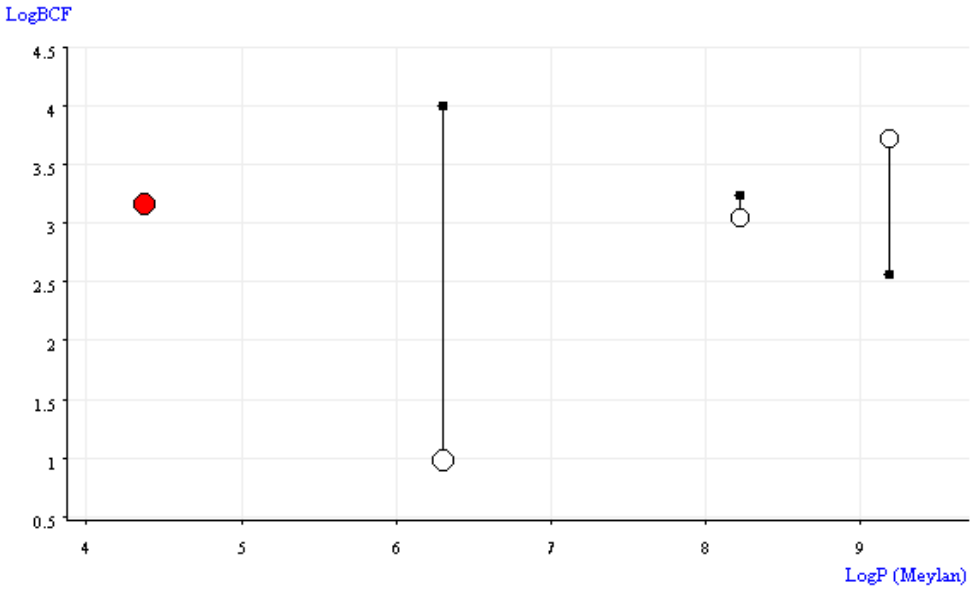
4.2 Reasoning:
Analysis of Molecular Descriptors

Descriptor name: LogP (Meylan)
Descriptor name: LogP is directly correlated to the logBCF value.

Following, a scatterplot of LogP (Meylan) against response values; experimental values are reported for the training set, predicted value for the studied compound. Light blue dots represent values of compounds from training set, red dot is the value of the studied compound.

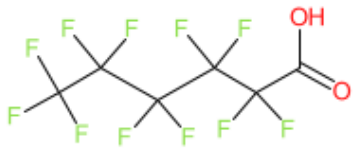


Following, a scatterplot of LogP (Meylan) against response values only for 3 most similar compounds in the training set. Red dot is the value of the studied compound, black outlined circles represents experimental values of compounds from training set, black dots represents predicted value of the same compound; the size of the circle is proportional to the similarity to the studied compound.



1. Prediction Summary

Prediction for compound Molecule 0 -



Prediction: Reliability:

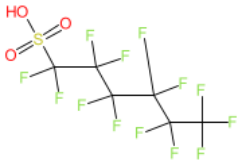
Prediction is 2.85 log(L/kg), but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Accuracy of prediction for similar molecules found in the training set is not optimal
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability

Compound: Molecule 0
Compound SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Experimental value: -
Predicted BCF [log(L/kg)]: 2.85
Molecules used for prediction: 4
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values

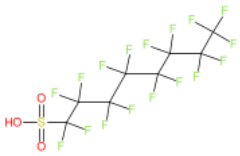
Compound #1

CAS: 335-67-1
Dataset id:308 (Training Set)
SMILES: O=C(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.906
Experimental value : 3.12
Predicted value : 3.096



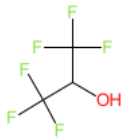
Compound #2

CAS: 355-46-4
Dataset id:310 (Training Set)
SMILES: O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.807
Experimental value : 3.6
Predicted value : 3.479



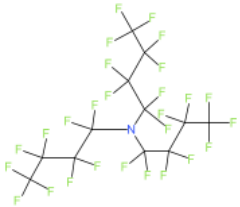
Compound #3

CAS: 1763-23-1
Dataset id:419 (Training Set)
SMILES: O=S(=O)(O)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.753
Experimental value : 3.73
Predicted value : 2.948



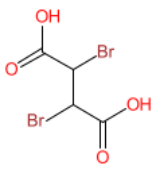
Compound #4

CAS: 920-66-1
Dataset id:391 (Training Set)
SMILES: FC(F)(F)C(O)C(F)(F)F
Similarity: 0.719
Experimental value : 0.244
Predicted value : 1.753



Compound #5

CAS: 311-89-7
Dataset id:303 (Training Set)
SMILES:
FC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)N(C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)F
Similarity: 0.649
Experimental value : 1.299
Predicted value : 3.432



Compound #6

CAS: 526-78-3
Dataset id:325 (Training Set)
SMILES: O=C(O)C(C(C(=O)O)Br)Br
Similarity: 0.618
Experimental value : 0.43
Predicted value : 0.02

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.7 Explanation: The predicted compound is outside the Applicability Domain of the model.
✔	Similar molecules with known experimental value Similarity index = 0.768 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.609 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 1.126 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✖	Maximum error of prediction among similar molecules Max error index = 1.509 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✔	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ The feature has a bad assessment, model is not reliable regarding this aspect.

References and Documentation

You can find complete details on each model and on how to read results in the proper model's guide, available on-line at www.vega-qsar.eu or directly in the VegaNIC application.

Mutagenicity (Ames test) CONSENSUS model(version 1.0.4)

Mutagenicity (Ames test) Consensus model based on the predictions of VEGA mutagenicity models.

Mutagenicity (Ames test) model (CAESAR)(version 2.1.14)

QSAR classification model for Mutagenicity (from CAESAR project)

Mutagenicity (Ames test) model (ISS)(version 1.0.3)

Classification model for Mutagenicity (Ames test) based on Benigni-Bossa (Istituto Superiore di Sanità) rule set

Mutagenicity (Ames test) model (SarPy-IRFMN)(version 1.0.8)

QSAR classification model for Mutagenicity (SarPy/IRFMN)

References and Documentation

Mutagenicity (Ames test) model (KNN-Read-Across)(version 1.0.1)

KNN (Read-Across) model for Mutagenicity (Ames test)

Mutagenicity (Ames test) model for aromatic amines (CONCERT/IRFMN)(version 1.0.0)

A knowledge-based expert rule system for predicting mutagenicity (AMES test) of aromatic amines.

Carcinogenicity model (CAESAR)(version 2.1.10)

QSAR classification model for Carcinogenicity (from CAESAR project)

Carcinogenicity model (ISS)(version 1.0.3)

Classification model for Carcinogenicity based on Benigni-Bossa (Istituto Superiore di Sanità) rule set

References and Documentation

Carcinogenicity model (IRFMN-ISSCAN-CGX)(version 1.0.2)

QSAR classification model for Carcinogenicity (IRFMN/ISSCAN-CGX) based on the ISSCAN-CGX dataset

Carcinogenicity model (IRFMN-Antares)(version 1.0.2)

QSAR classification model for Carcinogenicity (IRFMN/Antares) based on the Antares dataset

Carcinogenicity oral classification model (IRFMN)(version 1.0.1)

Classification model for carcinogenicity (oral route).

Carcinogenicity oral Slope Factor model (IRFMN)(version 1.0.1)

Quantitative model for the carcinogenicity (oral route) Slope Factor.

References and Documentation

Carcinogenicity inhalation classification model (IRFMN)(version 1.0.1)

Classification model for carcinogenicity (inhalation route).

Carcinogenicity inhalation Slope Factor model (IRFMN)(version 1.0.1)

Quantitative model for the carcinogenicity inhalation route) Slope Factor.

Carcinogenicity in male rat (CORAL)(version 1.0.0)

Carcinogenicity in male rat quantitative (TD50) model model (CORAL)

Carcinogenicity in female Rat (CORAL)(version 1.0.0)

Carcinogenicity in female rat quantitative (TD50) model model (CORAL)

References and Documentation

Acute Toxicity (LD50) model (KNN)(version 1.0.0)

KNN model for acute toxicity (LD50)

BCF model (CAESAR)(version 2.1.15)

QSAR regression model for fish BCF (from CAESAR project)

BCF model (Meylan)(version 1.0.4)

QSAR regression model for fish BCF (based on Meylan model)

BCF model (Arnot-Gobas)(version 1.0.1)

QSAR regression model for fish BCF (based on Arnot-Gobas model)

References and Documentation

BCF model (KNN-Read-Across)(version 1.1.1)

KNN (Read-Across) model for fish BCF