VEGA QSAR MODELS

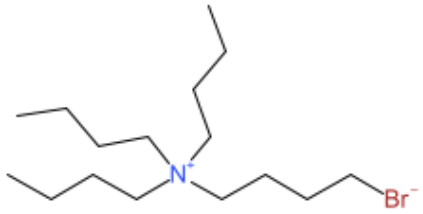
Report

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 -



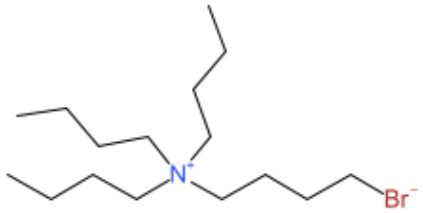
Prediction:

Prediction is Mutagenic with a consensus score of 0.15, based on 4 models.

Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Used models: 4
Predicted Consensus Mutagen activity: Mutagenic
Mutagenic Score: 0.15
Non-Mutagenic Score: 0.15
Model Caesar assessment: Suspect Mutagenic (LOW reliability)
Model ISS assessment: Mutagenic (LOW reliability)
Model SarPy assessment: Mutagenic (LOW reliability)
Model KNN assessment: NON-Mutagenic (MODERATE reliability)
Remarks:
none

1. Prediction Summary

Prediction for compound Molecule 0 -

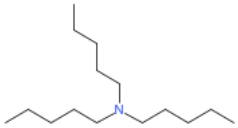


Prediction: Reliability:

Prediction is Suspect 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
- similar molecules found in the training set have experimental values that disagree with the predicted value
The following relevant fragments have been found: SA8 Aliphatic halogens

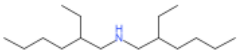
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Mutagen activity: Suspect Mutagenic
Structural Alerts: SA8 Aliphatic halogens
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



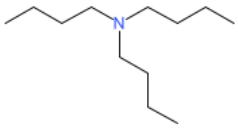
Compound #1

CAS: 621-77-2
Dataset id:349 (Training Set)
SMILES: N(CCCCC)(CCCCC)CCCCC
Similarity: 0.84
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #2

CAS: 106-20-7
Dataset id:357 (Training Set)
SMILES: N(CC(CC)CCCC)CC(CC)CCCC
Similarity: 0.832
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

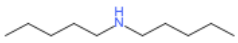


Compound #3

CAS: 102-82-9
Dataset id:2339 (Training Set)
SMILES: N(CCCC)(CCCC)CCCC
Similarity: 0.819
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

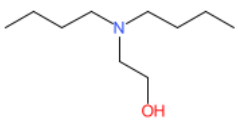
Compound #4

CAS: 143-16-8
Dataset id:638 (Training Set)
SMILES: N(CCCCCC)CCCCCC
Similarity: 0.798
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #5

CAS: 2050-92-2
Dataset id:1607 (Training Set)
SMILES: N(CCCCC)CCCCC
Similarity: 0.772
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #6

CAS: 102-81-8
Dataset id:2059 (Training Set)
SMILES: OCCN(CCCC)CCCC
Similarity: 0.756
Experimental value : NON-Mutagenic
Predicted value : NON-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 = 0.83 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 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 = 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.

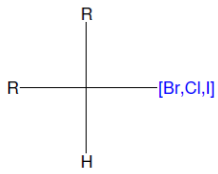
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ 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 :.

Fragment found: SA8 Aliphatic halogens



R = any atom/group

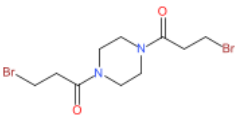
This alert contains non tertiary aliphatic halogens. Substances fired by Alerts SA2, SA4, SA5 and SA20 should be also excluded.

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

CAS: 112-52-7
Dataset id:3322 (Test Set)
SMILES: CCCCCCCCCCCCCl
Similarity: 0.687

Experimental value : NON-Mutagenic
Predicted value : Suspect Mutagenic

Alerts (found also in the target): SA8 Aliphatic halogens



CAS: 54-91-1
Dataset id:2154 (Training Set)
SMILES: O=C(N1CCN(C(=O)CCBr)CC1)CCBr
Similarity: 0.654

Experimental value : Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SA8 Aliphatic halogens

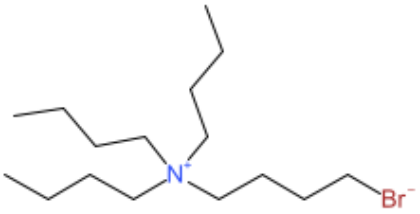
CAS: 630-17-1
Dataset id:2765 (Test Set)
SMILES: CC(C)(C)CBr
Similarity: 0.644

Experimental value : NON-Mutagenic
Predicted value : Suspect Mutagenic

Alerts (found also in the target): SA8 Aliphatic halogens

1. Prediction Summary

Prediction for compound Molecule 0 -

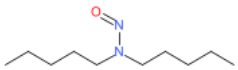


Prediction: Reliability:

Prediction is 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
- 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
- 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)
The following alerts have been found: SA8 Aliphatic halogens

Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Mutagen activity: Mutagenic
Structural Alerts: SA8 Aliphatic halogens
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

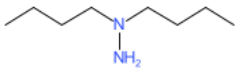
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 13256-06-9
Dataset id:886 (Training Set)
SMILES: O=NN(CCCCC)CCCCC
Similarity: 0.736
Experimental value : Mutagenic
Predicted value : Mutagenic

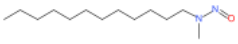
Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups



Compound #2

CAS: 7422-80-2
Dataset id:491 (Training Set)
SMILES: NN(CCCC)CCCC
Similarity: 0.735
Experimental value : NON-Mutagenic
Predicted value : Mutagenic

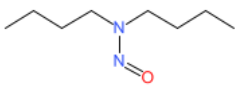
Alerts (not found also in the target): SA13 Hydrazine



Compound #3

CAS: 55090-44-3
Dataset id:547 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCC
Similarity: 0.72
Experimental value : Mutagenic
Predicted value : Mutagenic

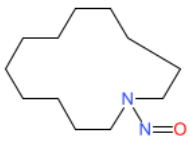
Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups



Compound #4

CAS: 924-16-3
Dataset id:515 (Training Set)
SMILES: O=NN(CCCC)CCCC
Similarity: 0.705
Experimental value : Mutagenic
Predicted value : Mutagenic

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

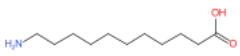


Compound #5

CAS: 40580-89-0
Dataset id:553 (Training Set)
SMILES: O=NN1CCCCCCCCCCC1
Similarity: 0.69
Experimental value : Mutagenic
Predicted value : Mutagenic

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

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



Compound #6

CAS: 2432-99-7
Dataset id:36 (Training Set)
SMILES: O=C(O)CCCCCCCCCCN
Similarity: 0.67
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.516 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.735 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.5 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 0.5 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 = 0.85 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.

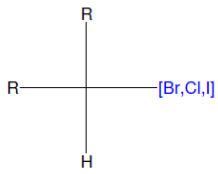
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ 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 fragments/structural alerts :.

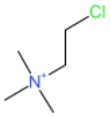
Fragment found: SA8 Aliphatic halogens



R = any atom/group

This alert contains non tertiary aliphatic halogens. Substances fired by Alerts SA2, SA4, SA5 and SA20 should be also excluded.

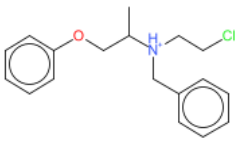
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 999-81-5
Dataset id:100 (Training Set)
SMILES: C[N+](C)(C)CCCl
Similarity: 0.634

Experimental value : NON-Mutagenic
Predicted value : Mutagenic

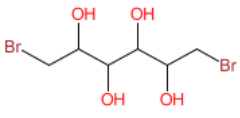
Alerts (found also in the target): SA8 Aliphatic halogens



CAS: 63-92-3
Dataset id:583 (Training Set)
SMILES: O(c1ccccc1)CC(C)[NH+](Cc2ccccc2)CCCl
Similarity: 0.589

Experimental value : Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SA8 Aliphatic halogens



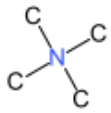
CAS: 10318-26-0
Dataset id:445 (Training Set)
SMILES: OC(CBr)C(O)C(O)C(O)CBr
Similarity: 0.586

Experimental value : Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SA8 Aliphatic halogens

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

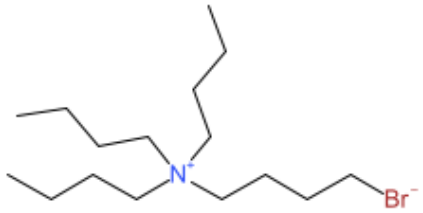
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



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

1. Prediction Summary

Prediction for compound Molecule 0 -

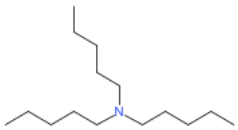


Prediction: Reliability:

Prediction is 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:
- similar molecules found in the training set have experimental values that disagree with the predicted value
The following relevant fragments have been found: SM93; SM142; SM175

Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Mutagen activity: Mutagenic
No. alerts for mutagenicity: 1
No. alerts for non-mutagenicity: 2
Structural Alerts: SM93; SM142; SM175
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

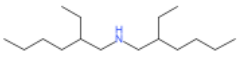
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 621-77-2
Dataset id:349 (Training Set)
SMILES: N(CCCCC)(CCCCC)CCCCC
Similarity: 0.84
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

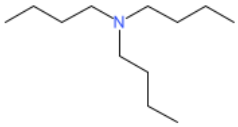
Alerts (found also in the target): SM142



Compound #2

CAS: 106-20-7
Dataset id:357 (Training Set)
SMILES: N(CC(CC)CCCC)CC(CC)CCCC
Similarity: 0.832
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (not found also in the target): SM163; SM169; SM182



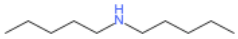
Compound #3

CAS: 102-82-9
Dataset id:2339 (Training Set)
SMILES: N(CCCC)(CCCC)CCCC
Similarity: 0.819
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM142

Compound #4

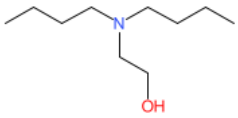
CAS: 143-16-8
Dataset id:638 (Training Set)
SMILES: N(CCCCCC)CCCCCC
Similarity: 0.798
Experimental value : NON-Mutagenic
Predicted value : Possible NON-Mutagenic



Compound #5

CAS: 2050-92-2
Dataset id:1607 (Training Set)
SMILES: N(CCCCC)CCCCC
Similarity: 0.772
Experimental value : NON-Mutagenic
Predicted value : Possible NON-Mutagenic

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



Compound #6

CAS: 102-81-8
Dataset id:2059 (Training Set)
SMILES: OCCN(CCCC)CCCC
Similarity: 0.756
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM142

Alerts (not found also in the target): SM128; SM176

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 = 0.83 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 = 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 = 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.

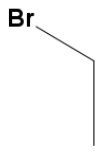
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

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

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

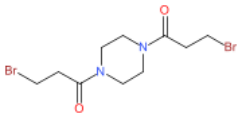
(Molecule 0) Reasoning on fragments/structural alerts - 1 of 3:.

Fragment found: SM93



Sarpy alert n. 93 for Mutagenicity, defined by SMARTS: C(C)Br

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



CAS: 54-91-1
Dataset id:2154 (Training Set)
SMILES: O=C(N1CCN(C(=O)CCBr)CC1)CCBr
Similarity: 0.654

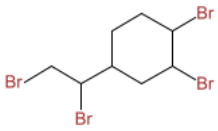
Experimental value : Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM93

CAS: 630-17-1
Dataset id:2765 (Test Set)
SMILES: CC(C)(C)CBr
Similarity: 0.644

Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM93



CAS: 3322-93-8
Dataset id:3607 (Test Set)
SMILES: C1CC(C(CC1C(CBr)Br)Br)Br
Similarity: 0.631

Experimental value : NON-Mutagenic
Predicted value : Mutagenic

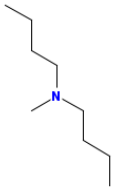
Alerts (found also in the target): SM93

Alerts (not found also in the target): SM77; SM163; SM169; SM182

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

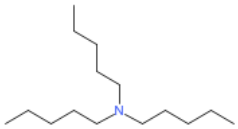
(Molecule 0) Reasoning on fragments/structural alerts - 2 of 3:.

Fragment found: SM142



Sarpy alert n. 142 for NON-Mutagenicity, defined by SMARTS: N(C)(CCCC)CCCC

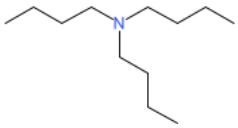
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 621-77-2
Dataset id:349 (Training Set)
SMILES: N(CCCCC)(CCCCC)CCCCC
Similarity: 0.84

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

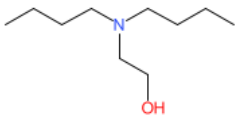
Alerts (found also in the target): SM142



CAS: 102-82-9
Dataset id:2339 (Training Set)
SMILES: N(CCCC)(CCCC)CCCC
Similarity: 0.819

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

Alerts (found also in the target): SM142



CAS: 102-81-8
Dataset id:2059 (Training Set)
SMILES: OCCN(CCCC)CCCC
Similarity: 0.756

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

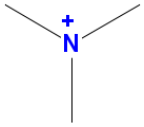
Alerts (found also in the target): SM142

Alerts (not found also in the target): SM128; SM176

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

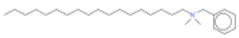
(Molecule 0) Reasoning on fragments/structural alerts - 3 of 3:.

Fragment found: SM175



Sarpy alert n. 175 for NON-Mutagenicity, defined by SMARTS: [N+](C)(C)C

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

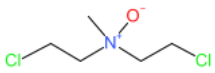


CAS: 37612-69-4
Dataset id:3982 (Training Set)
SMILES: c1ccc(cc1)C[N+](C)(C)CCCCCCCCCCCCCCCCCC
Similarity: 0.683

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

Alerts (found also in the target): SM175

Alerts (not found also in the target): SM157; SM163

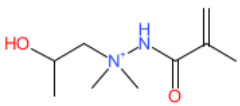


CAS: 126-85-2
Dataset id:4234 (Training Set)
SMILES: [O-][N+](C)(CCCl)CCCl
Similarity: 0.607

Experimental value : Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM175

Alerts (not found also in the target): SM73; SM106; SM204



CAS: 17341-40-1
Dataset id:66 (Training Set)
SMILES: O=C(N[N+](C)(C)CC(O)C)C(=C)C
Similarity: 0.602

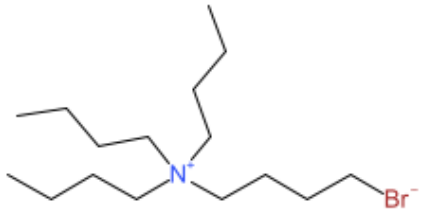
Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (found also in the target): SM175

Alerts (not found also in the target): SM103; SM176

1. Prediction Summary

Prediction for compound Molecule 0 -

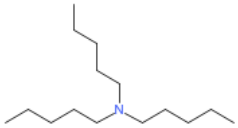


Prediction: Reliability:

Prediction is NON-Mutagenic, but the result shows some critical aspects, which require to be checked:

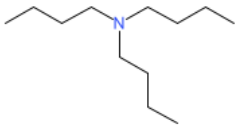
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
Molecules used for prediction: 4
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

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



Compound #1

CAS: 621-77-2
Dataset id:4102 (Training Set)
SMILES: N(CCCCC)(CCCCC)CCCCC
Similarity: 0.84
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

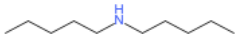


Compound #2

CAS: 102-82-9
Dataset id:119 (Training Set)
SMILES: N(CCCC)(CCCC)CCCC
Similarity: 0.819
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

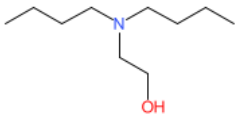
Compound #3

CAS: 143-16-8
Dataset id:1420 (Training Set)
SMILES: N(CCCCCC)CCCCCC
Similarity: 0.798
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



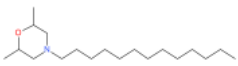
Compound #4

CAS: 2050-92-2
Dataset id:2133 (Training Set)
SMILES: N(CCCCC)CCCCC
Similarity: 0.772
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #5

CAS: 102-81-8
Dataset id:117 (Training Set)
SMILES: OCCN(CCCC)CCCC
Similarity: 0.756
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #6

CAS: 24602-86-6
Dataset id:2380 (Training Set)
SMILES: O1C(C)CN(CCCCCCCCCCCCC)CC1C
Similarity: 0.755
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

3.2 Applicability Domain:
Measured Applicability Domain Scores

⚠	Global AD Index AD index = 0.896 Explanation: The predicted compound could be out of the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.804 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..
✓	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:

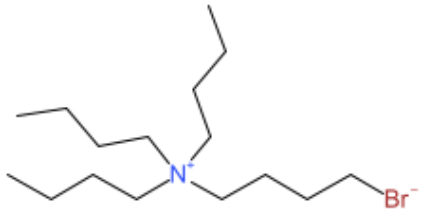
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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 -

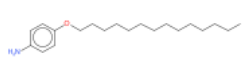


Prediction: Reliability:

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 (2 unknown fragments and 1 infrequent_fragments found)

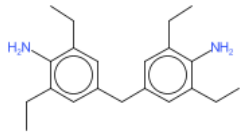
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Mutagen activity: NA
Structural Alerts: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



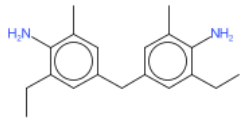
Compound #1

CAS: N.A.
Dataset id:6537 (Training Set)
SMILES: O(c1ccc(N)cc1)CCCCCCCCCCCCCC
Similarity: 0.648
Experimental value : NON-Mutagenic
Predicted value : Mutagenic



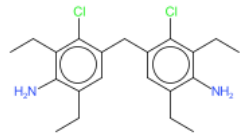
Compound #2

CAS: N.A.
Dataset id:5101 (Training Set)
SMILES: Nc1c(cc(cc1CC)Cc2cc(c(N)c(c2)CC)CC)CC
Similarity: 0.638
Experimental value : NON-Mutagenic
Predicted value : Mutagenic



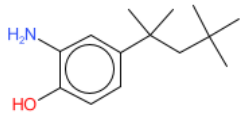
Compound #3

CAS: N.A.
Dataset id:5233 (Training Set)
SMILES: Nc1c(cc(cc1CC)Cc2cc(c(N)c(c2)CC)C)C
Similarity: 0.614
Experimental value : NON-Mutagenic
Predicted value : Mutagenic



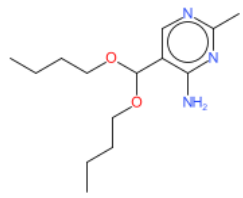
Compound #4

CAS: N.A.
Dataset id:6633 (Training Set)
SMILES: Nc1c(cc(c(c1CC)Cl)Cc2cc(c(N)c(c2Cl)CC)CC)CC
Similarity: 0.605
Experimental value : NON-Mutagenic
Predicted value : Mutagenic



Compound #5

CAS: N.A.
Dataset id:4177 (Training Set)
SMILES: Oc1ccc(cc1(N))C(C)(C)CC(C)(C)C
Similarity: 0.601
Experimental value : NON-Mutagenic
Predicted value : Mutagenic



Compound #6

CAS: N.A.
Dataset id:5176 (Training Set)
SMILES: n1cc(c(nc1C)N)C(OCCCC)OCCCC
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 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.643 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 = 0.34 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:

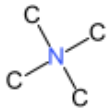
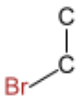
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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.

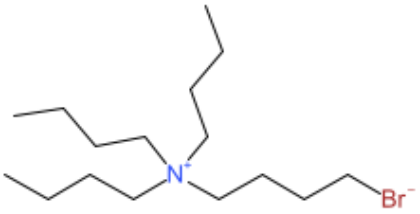
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

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

1. Prediction Summary

Prediction for compound Molecule 0 -

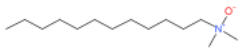


Prediction: Reliability:

Prediction is NON-Carcinogen, 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
- 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

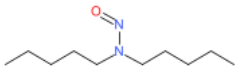
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Carcinogen activity: NON-Carcinogen
P(Carcinogen): 0.326
P(NON-Carcinogen): 0.674
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



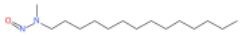
Compound #1

CAS: 1643-20-5
Dataset id:273 (Training Set)
SMILES: [O-][N+](C)(C)CCCCCCCCCCCC
Similarity: 0.788
Experimental value : NON-Carcinogen
Predicted value : Carcinogen



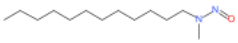
Compound #2

CAS: 13256-06-9
Dataset id:284 (Test Set)
SMILES: O=NN(CCCCC)CCCCC
Similarity: 0.736
Experimental value : Carcinogen
Predicted value : Carcinogen



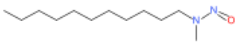
Compound #3

CAS: 75881-20-8
Dataset id:558 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCCCC
Similarity: 0.726
Experimental value : Carcinogen
Predicted value : Carcinogen



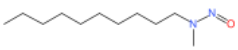
Compound #4

CAS: 55090-44-3
Dataset id:554 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCC
Similarity: 0.72
Experimental value : Carcinogen
Predicted value : Carcinogen



Compound #5

CAS: 68107-26-6
Dataset id:603 (Training Set)
SMILES: O=NN(C)CCCCCCCCCC
Similarity: 0.716
Experimental value : Carcinogen
Predicted value : Carcinogen



Compound #6

CAS: 75881-22-0
Dataset id:559 (Training Set)
SMILES: O=NN(C)CCCCCCCCCC
Similarity: 0.712
Experimental value : Carcinogen
Predicted value : Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.185 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.76 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✗	Accuracy of prediction for similar molecules Accuracy index = 0.477 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
⚠	Concordance for similar molecules Concordance index = 0.523 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 = 0.347 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:

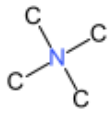
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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

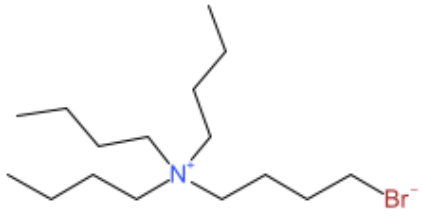
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



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

1. Prediction Summary

Prediction for compound Molecule 0 -

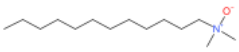


Prediction: Reliability:

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
- 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)
The following alerts have been found: SA8 Aliphatic halogens

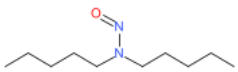
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Carcinogen activity: Carcinogen
Structural Alerts: SA8 Aliphatic halogens
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



Compound #1

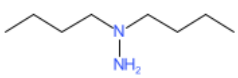
CAS: 1643-20-5
Dataset id:879 (Training Set)
SMILES: [O-][N+](C)(C)CCCCCCCCCCCC
Similarity: 0.788
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #2

CAS: 13256-06-9
Dataset id:886 (Training Set)
SMILES: O=NN(CCCCC)CCCCC
Similarity: 0.736
Experimental value : Carcinogen
Predicted value : Carcinogen

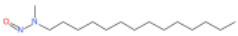
Alerts (not found also in the target): SA21 Alkyl and aryl N-nitroso groups



Compound #3

CAS: 7422-80-2
Dataset id:491 (Training Set)
SMILES: NN(CCCC)CCCC
Similarity: 0.735
Experimental value : Carcinogen
Predicted value : Carcinogen

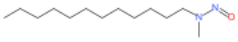
Alerts (not found also in the target): SA13 Hydrazine



Compound #4

CAS: 75881-20-8
Dataset id:579 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCCCC
Similarity: 0.726
Experimental value : Carcinogen
Predicted value : Carcinogen

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

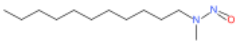


Compound #5

CAS: 55090-44-3
Dataset id:547 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCC
Similarity: 0.72
Experimental value : Carcinogen
Predicted value : Carcinogen

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

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



Compound #6

CAS: 68107-26-6
Dataset id:527 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCC
Similarity: 0.716
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 = 0.616 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.76 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 = 0.477 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
⚠	Atom Centered Fragments similarity check ACF index = 0.85 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.

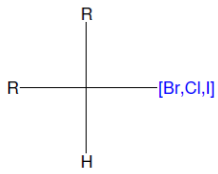
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

✖ 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 fragments/structural alerts :.

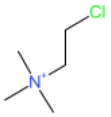
Fragment found: SA8 Aliphatic halogens



R = any atom/group

This alert contains non tertiary aliphatic halogens. Substances fired by Alerts SA2, SA4, SA5 and SA20 should be also excluded.

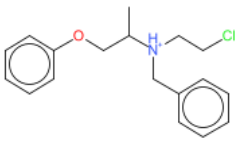
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 999-81-5
Dataset id:100 (Training Set)
SMILES: C[N+](C)(C)CCCl
Similarity: 0.634

Experimental value : NON-Carcinogen
Predicted value : Carcinogen

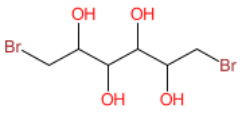
Alerts (found also in the target): SA8 Aliphatic halogens



CAS: 63-92-3
Dataset id:583 (Training Set)
SMILES: O(c1ccccc1)CC(C)[NH+](Cc2ccccc2)CCCl
Similarity: 0.589

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): SA8 Aliphatic halogens



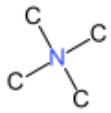
CAS: 10318-26-0
Dataset id:445 (Training Set)
SMILES: OC(CBr)C(O)C(O)C(O)CBr
Similarity: 0.586

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): SA8 Aliphatic halogens

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

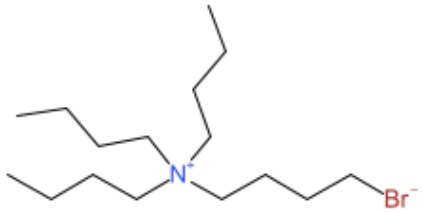
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



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

1. Prediction Summary

Prediction for compound Molecule 0 -

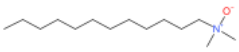


Prediction: Reliability:

Prediction is Carcinogen, but the result shows some critical aspects, which require to be checked:
- 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
The following relevant fragments have been found: Carcinogenity alert no. 4

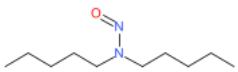
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Carcinogenic activity: Carcinogen
No. alerts for carcinogenicity: 1
Structural Alerts: Carcinogenity alert no. 4
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

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



Compound #1

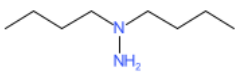
CAS: 1643-20-5
Dataset id:777 (Training Set)
SMILES: [O-][N+](C)(C)CCCCCCCCCCCC
Similarity: 0.788
Experimental value : NON-Carcinogen
Predicted value : Possible NON-Carcinogen



Compound #2

CAS: 13256-06-9
Dataset id:638 (Training Set)
SMILES: O=NN(CCCCC)CCCCC
Similarity: 0.736
Experimental value : Carcinogen
Predicted value : Carcinogen

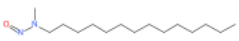
Alerts (not found also in the target): Carcinogenity alert no. 1; Carcinogenity alert no. 14; Carcinogenity alert no. 27



Compound #3

CAS: 7422-80-2
Dataset id:406 (Training Set)
SMILES: NN(CCCC)CCCC
Similarity: 0.735
Experimental value : Carcinogen
Predicted value : Carcinogen

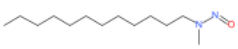
Alerts (not found also in the target): Carcinogenity alert no. 27



Compound #4

CAS: 75881-20-8
Dataset id:489 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCCCC
Similarity: 0.726
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

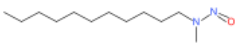


Compound #5

CAS: 55090-44-3
Dataset id:458 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCC
Similarity: 0.72
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

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



Compound #6

CAS: 68107-26-6
Dataset id:439 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCC
Similarity: 0.716
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

3.2 Applicability Domain:
Measured Applicability Domain Scores

⚠	Global AD Index AD index = 0.777 Explanation: The predicted compound could be out of the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.751 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 = 0.645 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.

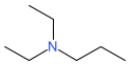
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

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

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

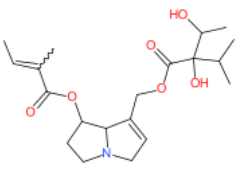
(Molecule 0) Reasoning on fragments/structural alerts :.

Fragment found: Carcinogenity alert no. 4



Structural alert for carcinogenity defined by the SMARTS:CCCN(CC)CC

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

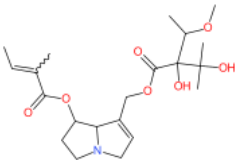


CAS: 22571-95-5
Dataset id:403 (Training Set)
SMILES: O=C(OC2CCN1CC=C(COC(=O)C(O)(C(O)C)C(C)C)C12)C(=CC)C
Similarity: 0.567

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 4

Alerts (not found also in the target): Carcinogenity alert no. 20; Carcinogenity alert no. 29

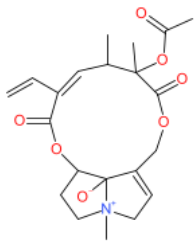


CAS: 303-34-4
Dataset id:160 (Training Set)
SMILES: O=C(OC2CCN1CC=C(COC(=O)C(O)(C(OC)C)C(O)(C)C)C12)C(=CC)C
Similarity: 0.553

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 4

Alerts (not found also in the target): Carcinogenity alert no. 20; Carcinogenity alert no. 29



CAS: 33979-15-6
Dataset id:384 (Training Set)
SMILES: O=C1OC2CC[N+]3(C)(CC=C(COC(=O)C(OC(=O)C)(C)C(C=C1(C=C))C)C23([O-]))
Similarity: 0.552

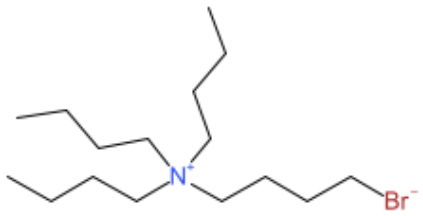
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 4

Alerts (not found also in the target): Carcinogenity alert no. 13; Carcinogenity alert no. 20; Carcinogenity alert no. 29

1. Prediction Summary

Prediction for compound Molecule 0 -

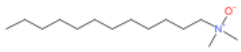


Prediction: Reliability:

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 relevant fragments have been found: Carcinogenity alert no. 58; Carcinogenity alert no. 59

Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Carcinogenic activity: Carcinogen
No. alerts for carcinogenicity: 2
Structural Alerts: Carcinogenity alert no. 58; Carcinogenity alert no. 59
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

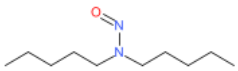
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 1643-20-5
Dataset id:273 (Training Set)
SMILES: [O-][N+](C)(C)CCCCCCCCCCCC
Similarity: 0.788
Experimental value : NON-Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 64



Compound #2

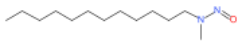
CAS: 13256-06-9
Dataset id:284 (Test Set)
SMILES: O=NN(CCCCC)CCCCC
Similarity: 0.736
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 5; Carcinogenity alert no. 8; Carcinogenity alert no. 9; 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 #3

CAS: 75881-20-8
Dataset id:558 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCCCC
Similarity: 0.726
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 4; Carcinogenity alert no. 8; Carcinogenity alert no. 9; Carcinogenity alert no. 10; Carcinogenity alert no. 15; Carcinogenity alert no. 50; Carcinogenity alert no. 51; Carcinogenity alert no. 54; Carcinogenity alert no. 55; Carcinogenity alert no. 63

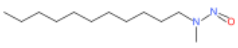


Compound #4

CAS: 55090-44-3
Dataset id:554 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCCC
Similarity: 0.72
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 4; Carcinogenity alert no. 8; Carcinogenity alert no. 9; Carcinogenity alert no. 10; Carcinogenity alert no. 15; Carcinogenity alert no. 50; Carcinogenity alert no. 51; Carcinogenity alert no. 54; Carcinogenity alert no. 55; Carcinogenity alert no. 63

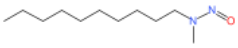
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #5

CAS: 68107-26-6
Dataset id:603 (Training Set)
SMILES: O=NN(C)CCCCCCCCCCC
Similarity: 0.716
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 4; Carcinogenity alert no. 8; Carcinogenity alert no. 9; Carcinogenity alert no. 10; Carcinogenity alert no. 15; Carcinogenity alert no. 50; Carcinogenity alert no. 51; Carcinogenity alert no. 54; Carcinogenity alert no. 55; Carcinogenity alert no. 63



Compound #6

CAS: 75881-22-0
Dataset id:559 (Training Set)
SMILES: O=NN(C)CCCCCCCCCC
Similarity: 0.712
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 4; Carcinogenity alert no. 8; Carcinogenity alert no. 9; Carcinogenity alert no. 10; Carcinogenity alert no. 15; Carcinogenity alert no. 50; Carcinogenity alert no. 51; Carcinogenity alert no. 54; Carcinogenity alert no. 55; Carcinogenity alert no. 63

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.416 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.747 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.643 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 0.643 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.

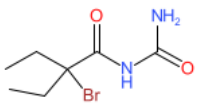
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

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

Fragment found: Carcinogenity alert no. 58

Structural alert for carcinogenity defined by the SMARTS: CCBr

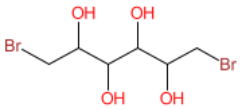
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 77-65-6
Dataset id:129 (Training Set)
SMILES: O=C(N)NC(=O)C(CC)(CC)Br
Similarity: 0.619

Experimental value : NON-Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 58; Carcinogenity alert no. 59



CAS: 10318-26-0
Dataset id:213 (Training Set)
SMILES: OC(CBr)C(O)C(O)C(O)CBr
Similarity: 0.586

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 58; Carcinogenity alert no. 59

CAS: 3296-90-0
Dataset id:93 (Training Set)
SMILES: OCC(CO)(CBr)CBr
Similarity: 0.58

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 58; Carcinogenity alert no. 59

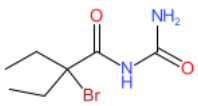
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

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

Fragment found: Carcinogenity alert no. 59

Structural alert for carcinogenity defined by the SMARTS: CBr

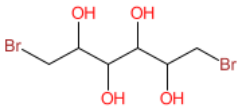
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 77-65-6
Dataset id:129 (Training Set)
SMILES: O=C(N)NC(=O)C(CC)(CC)Br
Similarity: 0.619

Experimental value : NON-Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 58; Carcinogenity alert no. 59



CAS: 10318-26-0
Dataset id:213 (Training Set)
SMILES: OC(CBr)C(O)C(O)C(O)CBr
Similarity: 0.586

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 58; Carcinogenity alert no. 59

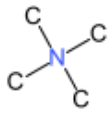
CAS: 3296-90-0
Dataset id:93 (Training Set)
SMILES: OCC(CO)(CBr)CBr
Similarity: 0.58

Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (found also in the target): Carcinogenity alert no. 58; Carcinogenity alert no. 59

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

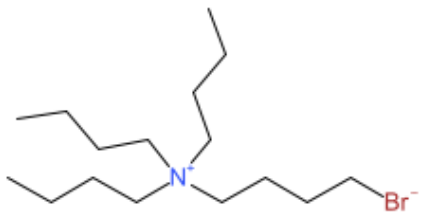
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



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

1. Prediction Summary

Prediction for compound Molecule 0 -

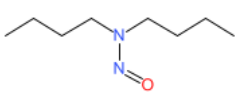


Prediction: Reliability:

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
- 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)

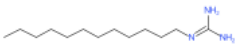
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Oral Carcinogenic class: Carcinogen
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



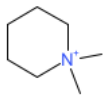
Compound #1

CAS: 924-16-3
Dataset id:224 (Test Set)
SMILES: O=NN(CCCC)CCCC
Similarity: 0.705
Experimental value : Carcinogen
Predicted value : Carcinogen



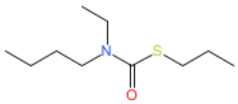
Compound #2

CAS: 2439-10-3
Dataset id:490 (Training Set)
SMILES: N(=C(N)N)CCCCCCCCCCCC
Similarity: 0.685
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



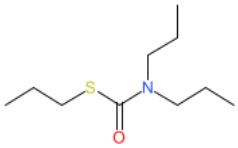
Compound #3

CAS: 24307-26-4
Dataset id:566 (Training Set)
SMILES: C[N+]1(C)(CCCCC1)
Similarity: 0.682
Experimental value : NON-Carcinogen
Predicted value : Carcinogen



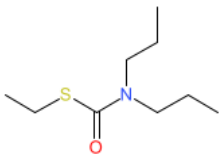
Compound #4

CAS: 1114-71-2
Dataset id:622 (Training Set)
SMILES: O=C(N(CC)CCCC)SCCC
Similarity: 0.682
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #5

CAS: 1929-77-7
Dataset id:738 (Test Set)
SMILES: O=C(N(CCC)CCC)SCCC
Similarity: 0.673
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #6

CAS: 759-94-4
Dataset id:491 (Training Set)
SMILES: O=C(N(CCC)CCC)SCC
Similarity: 0.657
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.599 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.695 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 = 0.511 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:

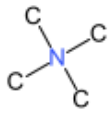
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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

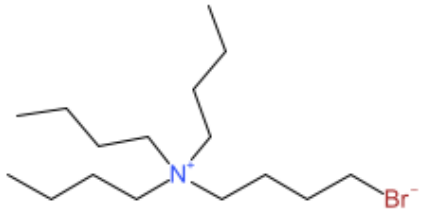
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



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

1. Prediction Summary

Prediction for compound Molecule 0 -

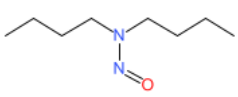


Prediction: Reliability:

Prediction is 0.19, 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 (1 unknown fragments found)

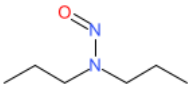
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Oral Carcinogenicity SF (log form) [log(1/(mg/kg-day))]: 0.19
Predicted Oral Carcinogenicity SF [1/(mg/kg-day)]: 1.55
Experimental value [1/(mg/kg-day)]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



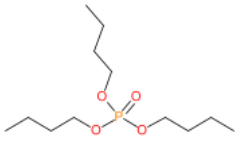
Compound #1

CAS: 924-16-3
Dataset id:224 (Test Set)
SMILES: O=NN(CCCC)CCCC
Similarity: 0.705
Experimental value : 0.73
Predicted value : 0.473



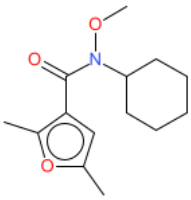
Compound #2

CAS: 621-64-7
Dataset id:225 (Training Set)
SMILES: O=NN(CCC)CCC
Similarity: 0.626
Experimental value : 0.85
Predicted value : 0.936



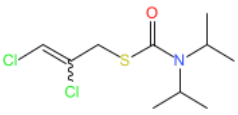
Compound #3

CAS: 126-73-8
Dataset id:299 (Training Set)
SMILES: O=P(OCCCC)(OCCCC)OCCCC
Similarity: 0.626
Experimental value : -2.05
Predicted value : -2.145



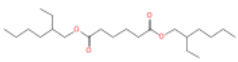
Compound #4

CAS: 60568-05-0
Dataset id:151 (Training Set)
SMILES: O=C(c1cc(oc1C)C)N(OC)C2CCCCC2
Similarity: 0.623
Experimental value : -1.52
Predicted value : -1.408



Compound #5

CAS: 2303-16-4
Dataset id:95 (Training Set)
SMILES: O=C(N(C(C)C)C(C)C)SCC(=CCl)Cl
Similarity: 0.615
Experimental value : -1.21
Predicted value : -1.799



Compound #6

CAS: 103-23-1
Dataset id:94 (Test Set)
SMILES: O=C(OCC(CC)CCCC)CCCCC(=O)OCC(CC)CCCC
Similarity: 0.614
Experimental value : -2.92
Predicted value : -1.999

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.397 Explanation: The predicted compound is outside the Applicability Domain of the model.
✗	Similar molecules with known experimental value Similarity index = 0.661 Explanation: No similar compounds with known experimental value in the training set have been found..
✓	Accuracy of prediction for similar molecules Accuracy index = 0.171 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✓	Concordance for similar molecules Concordance index = 0.6 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 = 0.257 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 = 0.6 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:

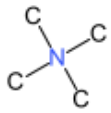
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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

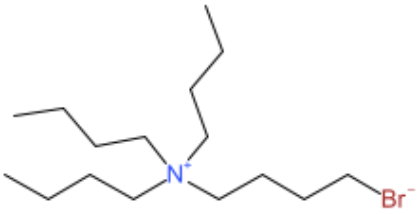
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



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

1. Prediction Summary

Prediction for compound Molecule 0 -

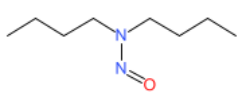


Prediction: Reliability:

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)

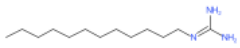
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Inhalation Carcinogenic class: Carcinogen
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



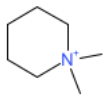
Compound #1

CAS: 924-16-3
Dataset id:192 (Training Set)
SMILES: O=NN(CCCC)CCCC
Similarity: 0.705
Experimental value : Carcinogen
Predicted value : Carcinogen



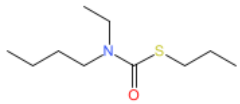
Compound #2

CAS: 2439-10-3
Dataset id:462 (Training Set)
SMILES: N(=C(N)N)CCCCCCCCCCCC
Similarity: 0.685
Experimental value : NON-Carcinogen
Predicted value : Carcinogen



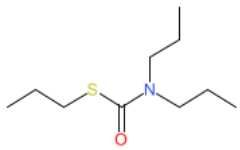
Compound #3

CAS: 24307-26-4
Dataset id:545 (Test Set)
SMILES: C[N+]1(C)(CCCCC1)
Similarity: 0.682
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



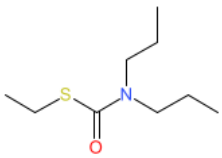
Compound #4

CAS: 1114-71-2
Dataset id:605 (Test Set)
SMILES: O=C(N(CC)CCCC)SCCC
Similarity: 0.682
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #5

CAS: 1929-77-7
Dataset id:743 (Training Set)
SMILES: O=C(N(CCC)CCC)SCCC
Similarity: 0.673
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #6

CAS: 759-94-4
Dataset id:463 (Training Set)
SMILES: O=C(N(CCC)CCC)SCC
Similarity: 0.657
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.358 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.695 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✖	Accuracy of prediction for similar molecules Accuracy index = 0.511 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✖	Concordance for similar molecules Concordance index = 0.511 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 = 0.6 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:

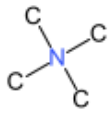
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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

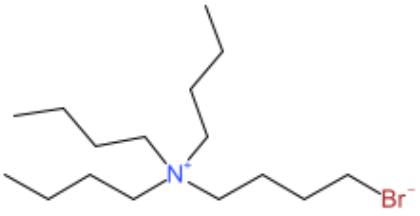
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



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

1. Prediction Summary

Prediction for compound Molecule 0 -

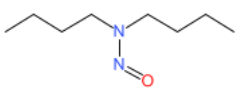


Prediction: Reliability:

Prediction is 0.3, 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
- the maximum error in prediction of similar molecules found in the training set has a moderate 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 (1 unknown fragments found)

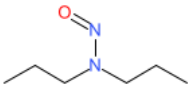
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted Inhalation Carcinogenicity SF (log form) [log(1/(mg/kg-day))]: 0.3
Predicted Inhalation Carcinogenicity SF [1/(mg/kg-day)]: 1.97
Experimental value [1/(mg/kg-day)]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



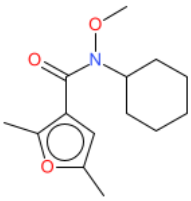
Compound #1

CAS: 924-16-3
Dataset id:192 (Training Set)
SMILES: O=NN(CCCC)CCCC
Similarity: 0.705
Experimental value : 0.75
Predicted value : -0.335



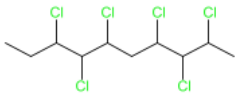
Compound #2

CAS: 621-64-7
Dataset id:193 (Training Set)
SMILES: O=NN(CCC)CCC
Similarity: 0.626
Experimental value : 0.85
Predicted value : -0.147



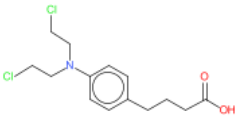
Compound #3

CAS: 60568-05-0
Dataset id:123 (Training Set)
SMILES: O=C(c1cc(oc1C)C)N(OC)C2CCCCC2
Similarity: 0.623
Experimental value : -1.52
Predicted value : 0.268



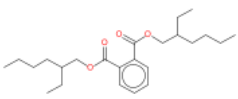
Compound #4

CAS: 108171-26-2
Dataset id:54 (Training Set)
SMILES: CCC(C(C(CC(C(C(C)Cl)Cl)Cl)Cl)Cl)Cl
Similarity: 0.602
Experimental value : -1.06
Predicted value : 0.58



Compound #5

CAS: 305-03-3
Dataset id:50 (Training Set)
SMILES: O=C(O)CCCc1ccc(cc1)N(CCCl)CCCl
Similarity: 0.591
Experimental value : 2.66
Predicted value : 1.001



Compound #6

CAS: 117-81-7
Dataset id:38 (Training Set)
SMILES: O=C(OCC(CC)CCCC)c1ccccc1(C(=O)OCC(CC)CCCC)
Similarity: 0.589
Experimental value : -2.08
Predicted value : 0.163

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.397 Explanation: The predicted compound is outside the Applicability Domain of the model.
✖	Similar molecules with known experimental value Similarity index = 0.661 Explanation: No similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 1.041 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✔	Concordance for similar molecules Concordance index = 0.505 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 = 1.085 Explanation: the maximum error in prediction of similar molecules found in the training set has a moderate 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 = 0.6 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:

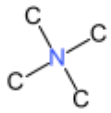
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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

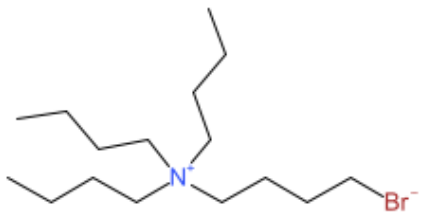
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



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

1. Prediction Summary

Prediction for compound Molecule 0 -

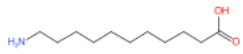


Prediction: Reliability:

Prediction is -1.0975, 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
- 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)

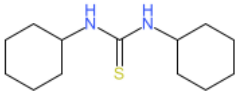
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted TD50 [-log(mg/kg bw/day)]: -1.0975
Predicted TD50 [mg/kg bw/day]: 12.52
Experimental TD50 [mg/kg bw/day]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



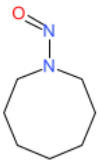
Compound #1

CAS: N.A.
Dataset id:34 (Training Set)
SMILES: C(CCCCCCCCCC(=O)O)N
Similarity: 0.67
Experimental value : -3.041
Predicted value : -2.357



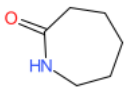
Compound #2

CAS: N.A.
Dataset id:176 (Test Set)
SMILES: C1C(CCCC1)NC(=S)NC1CCCCC1
Similarity: 0.664
Experimental value : -4.193
Predicted value : -3.794



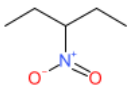
Compound #3

CAS: N.A.
Dataset id:185 (Test Set)
SMILES: C1CCCCCCN1N=O
Similarity: 0.628
Experimental value : 0.703
Predicted value : -0.162



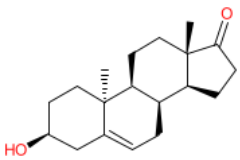
Compound #4

CAS: N.A.
Dataset id:141 (Training Set)
SMILES: C1CCNC(=O)CC1
Similarity: 0.615
Experimental value : -3.755
Predicted value : -2.929



Compound #5

CAS: N.A.
Dataset id:5 (Training Set)
SMILES: C(CC)(CC)N(=O)=O
Similarity: 0.599
Experimental value : -0.375
Predicted value : -1.209



Compound #6

CAS: N.A.
Dataset id:164 (Test Set)
SMILES:
C12=CC[C@H]3[C@@H]([C@@]2(CC[C@@H](C1)O)C)CC[C@]1([C@H]3CCC1=O)C
Similarity: 0.597
Experimental value : -1.833
Predicted value : -0.788

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.4 Explanation: The predicted compound is outside the Applicability Domain of the model.
✗	Similar molecules with known experimental value Similarity index = 0.667 Explanation: No similar compounds with known experimental value in the training set have been found..
✓	Accuracy of prediction for similar molecules Accuracy index = 0.542 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✗	Concordance for similar molecules Concordance index = 2.519 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 = 0.684 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 = 0.6 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:

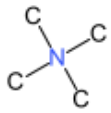
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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

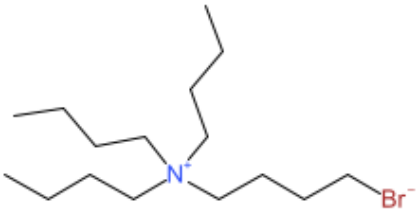
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



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

1. Prediction Summary

Prediction for compound Molecule 0 -

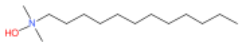


Prediction: Reliability:

Prediction is 0.0753, 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
- 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 (1 unknown fragments and 1 infrequent_fragments found)

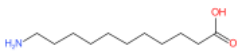
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted TD50 [-log(mg/kg bw/day)]: 0.0753
Predicted TD50 [mg/kg bw/day]: 0.8408
Experimental TD50 [mg/kg bw/day]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



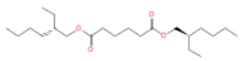
Compound #1

CAS: N.A.
Dataset id:93 (Training Set)
SMILES: CCCCCCCCCCCC[N](O)(C)C
Similarity: 0.774
Experimental value : -2.364
Predicted value : -0.843



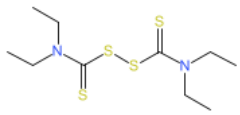
Compound #2

CAS: N.A.
Dataset id:117 (Training Set)
SMILES: C(CCCCCCCCCC(=O)O)N
Similarity: 0.67
Experimental value : -4.649
Predicted value : -4.782



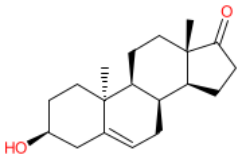
Compound #3

CAS: N.A.
Dataset id:113 (Training Set)
SMILES: CCCC[C@H](CC)COC(=O)CCCCC(=O)OC[C@H](CC)CCCC
Similarity: 0.614
Experimental value : -4.161
Predicted value : -2.71



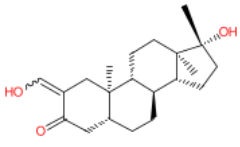
Compound #4

CAS: N.A.
Dataset id:102 (Training Set)
SMILES: N(CC)(CC)C(=S)SSC(=S)N(CC)CC
Similarity: 0.602
Experimental value : -3.155
Predicted value : -3.184



Compound #5

CAS: N.A.
Dataset id:25 (Training Set)
SMILES:
C12=CC[C@H]3[C@@H]([C@@]2(CC[C@@H](C1)O)C)CC[C@]1([C@H]3CCC1=O)C
Similarity: 0.597
Experimental value : -1.921
Predicted value : -4.971



Compound #6

CAS: N.A.
Dataset id:31 (Training Set)
SMILES:
C1(=O)C[C@@H]2[C@@](C/C/1=C\O)([C@H]1[C@H](CC2)[C@@H]2[C@@](CC1)([C@@](CC2)(O)C)C)C
Similarity: 0.591
Experimental value : -2.279
Predicted value : -3.484

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.364 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.714 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.827 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✗	Concordance for similar molecules Concordance index = 3.582 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 = 1.521 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 = 0.51 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:

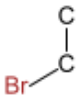
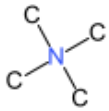
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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.

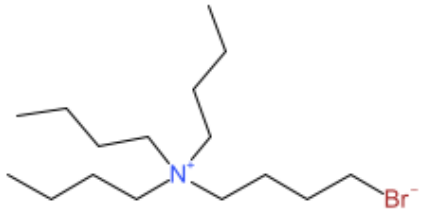
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

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

1. Prediction Summary

Prediction for compound Molecule 0 -

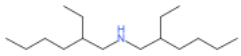


Prediction: Reliability:

Prediction is 4317.92 mg/kg, 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
- the maximum error in prediction of similar molecules found in the training set has a moderate value, considering the experimental variability

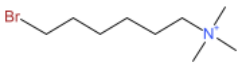
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted log LD50 [log(mmol/Kg)]: 1.128
Predicted log LD50 [mg/Kg]: 4317.92
Molecules used for prediction: 3
Experimental value [mg/Kg]: -
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

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



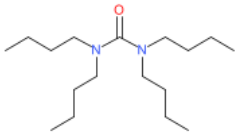
Compound #1

CAS: N.A.
Dataset id:1890 (Training Set)
SMILES: N(CC(CC)CCCC)CC(CC)CCCC
Similarity: 0.832
Experimental value : 0.83
Predicted value : 0.42



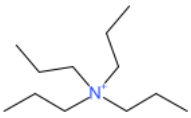
Compound #2

CAS: N.A.
Dataset id:4907 (Training Set)
SMILES: C[N+](C)(C)CCCCCCBr
Similarity: 0.813
Experimental value : 0.82
Predicted value : 0.235



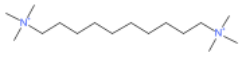
Compound #3

CAS: N.A.
Dataset id:2047 (Training Set)
SMILES: O=C(N(CCCC)CCCC)N(CCCC)CCCC
Similarity: 0.807
Experimental value : 1.77
Predicted value : 0.596



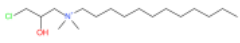
Compound #4

CAS: N.A.
Dataset id:995 (Training Set)
SMILES: CCC[N+](CCC)(CCC)CCC
Similarity: 0.806
Experimental value : 1.1
Predicted value : 0.492



Compound #5

CAS: N.A.
Dataset id:1895 (Training Set)
SMILES: C[N+](C)(C)CCCCCCCCCC[N+](C)(C)C
Similarity: 0.804
Experimental value : -0.78
Predicted value : 0.517



Compound #6

CAS: N.A.
Dataset id:2049 (Training Set)
SMILES: OC(C[N+](C)(C)CCCCCCCCCCCC)CCl
Similarity: 0.794
Experimental value : 0.5
Predicted value : 0.429

3.2 Applicability Domain:
Measured Applicability Domain Scores

⚠	Global AD Index AD index = 0.85 Explanation: The predicted compound could be out of the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.817 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.723 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✓	Concordance for similar molecules Concordance index = 0.416 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 = 1.174 Explanation: the maximum error in prediction of similar molecules found in the training set has a moderate 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:

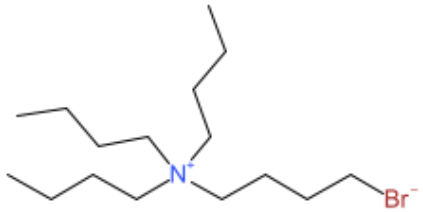
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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 0.41 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 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
- 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)
The following relevant fragments have been found: Tertiary amine (SR 05)

Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted BCF [log(L/kg)]: 0.41
Predicted BCF [L/kg]: 3
Predicted BCF from sub-model 1 (HM) [log(L/kg)]: 0.83
Predicted BCF from sub-model 2 (GA) [log(L/kg)]: 0.31
Predicted LogP (MLogP): 1.19
Structural Alerts: Tertiary amine (SR 05)
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

2. Possible Use and Uncertainty

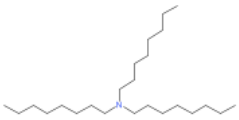
Threshold 3.3 (bioaccumulative)

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.

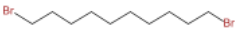
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

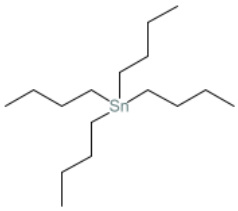
CAS: 1116-76-3
Dataset id:306 (Training Set)
SMILES: N(CCCCCCCC)(CCCCCCCC)CCCCCCCC
Similarity: 0.757
Experimental value : 1.92
Predicted value : 1.35

Alerts (found also in the target): Tertiary amine (SR 05)



Compound #2

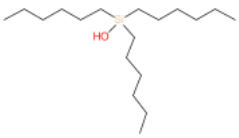
CAS: 4101-68-2
Dataset id:251 (Test Set)
SMILES: C(CCCCCBr)CCCCBr
Similarity: 0.756
Experimental value : 1.78
Predicted value : 2.897



Compound #3

CAS: 1461-25-2
Dataset id:467 (Training Set)
SMILES: CCCC[Sn](CCCC)(CCCC)CCCC
Similarity: 0.7
Experimental value : 2.34
Predicted value : 2.987

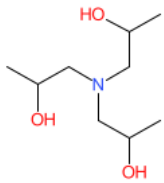
Alerts (not found also in the target): Sn atom in the molecule (SO 04)



Compound #4

CAS: 60782-58-3
Dataset id:465 (Training Set)
SMILES: O[Si](CCCCCC)(CCCCCC)CCCCCC
Similarity: 0.685
Experimental value : 1.48
Predicted value : 1.927

Alerts (not found also in the target): Si atom in the molecule (SO 03); OH group (PG 06)



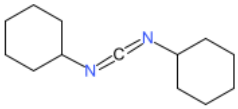
Compound #5

CAS: 122-20-3
Dataset id:311 (Training Set)
SMILES: OC(C)CN(CC(O)C)CC(O)C
Similarity: 0.654
Experimental value : -0.24
Predicted value : 0.004

Alerts (found also in the target): Tertiary amine (SR 05)

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

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



Compound #6

CAS: 538-75-0
Dataset id:373 (Training Set)
SMILES: C(=NC1CCCCC1)=NC2CCCCC2
Similarity: 0.653
Experimental value : 0.34
Predicted value : 1.39

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.643 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.756 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.844 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✖	Concordance for similar molecules Concordance index = 1.442 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 = 1.117 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 = 0.85 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.

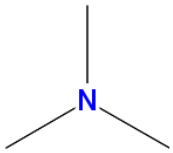
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

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

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

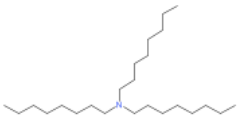
(Molecule 0) Reasoning on fragments/structural alerts :.

Fragment found: Tertiary amine (SR 05)



This chemical contains a tertiary amine. This residue has been found to be present in a large numberof non-bioaccumulative compounds (28), even when the logP value was higher than 3.

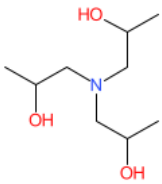
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 1116-76-3
Dataset id:306 (Training Set)
SMILES: N(CCCCCCCC)(CCCCCCCC)CCCCCCCC
Similarity: 0.757

Experimental value : 1.92
Predicted value : 1.35

Alerts (found also in the target): Tertiary amine (SR 05)

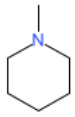


CAS: 122-20-3
Dataset id:311 (Training Set)
SMILES: OC(C)CN(CC(O)C)CC(O)C
Similarity: 0.654

Experimental value : -0.24
Predicted value : 0.004

Alerts (found also in the target): Tertiary amine (SR 05)

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



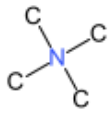
CAS: 626-67-5
Dataset id:441 (Training Set)
SMILES: N1(C)CCCCC1
Similarity: 0.617

Experimental value : 0.65
Predicted value : 0.459

Alerts (found also in the target): Tertiary amine (SR 05)

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:

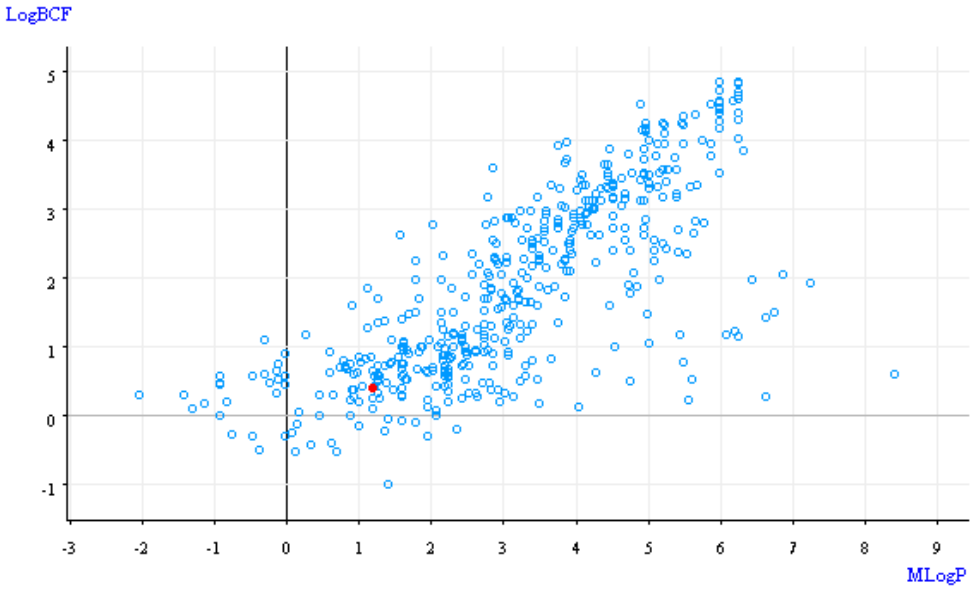


Fragment defined by the SMILES: CN(C)(C)C
The fragment has less than 3 occurrences in the model's training set

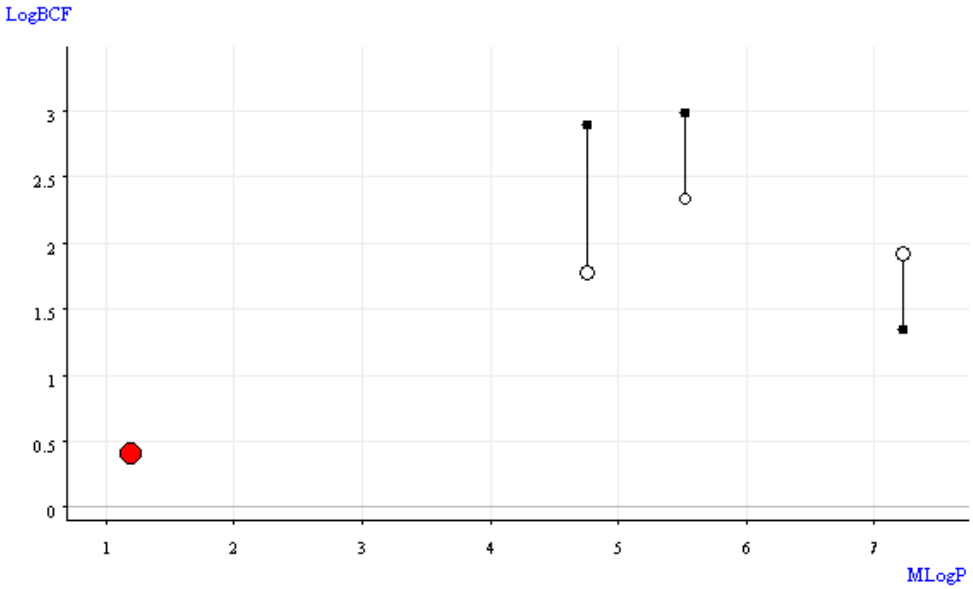
4.2 Reasoning:
Analysis of Molecular Descriptors

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.

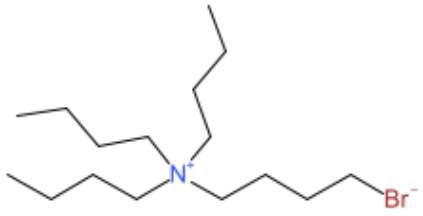


Following, a scatterplot of MLogP against response values only for 3 most similar compounds in the training set. Red dot is the value of the studied compound, black outlined circles represents experimental values of compounds from training set, black dots represents predicted value of the same compound;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



1. Prediction Summary

Prediction for compound Molecule 0 -

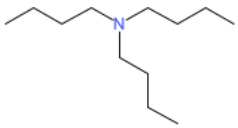


Prediction: Reliability:

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:
- Only moderately similar compounds with known experimental value in the training set have been found
- 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 moderate value, considering the experimental variability
- 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 found)

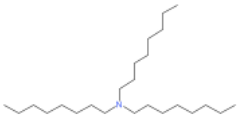
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted BCF [log(L/kg)]: 0.5
Predicted BCF [L/kg]: 3
Predicted LogP (Meylan/Kowwin): 2.05
Predicted LogP reliability: Low
MW: 319.73
Ionic compound: yes
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



Compound #1

CAS: 102-82-9
Dataset id:614 (Test Set)
SMILES: N(CCCC)(CCCC)CCCC
Similarity: 0.819
Experimental value : 1.4
Predicted value : 1.234

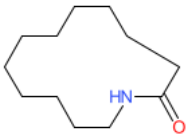


Compound #2

CAS: 1116-76-3
Dataset id:647 (Test Set)
SMILES: N(CCCCCCCC)(CCCCCCCC)CCCCCCCC
Similarity: 0.757
Experimental value : 1.93
Predicted value : 2.482

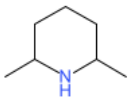
Compound #3

CAS: 4101-68-2
Dataset id:459 (Training Set)
SMILES: C(CCCCCBr)CCCCBr
Similarity: 0.756
Experimental value : 1.78
Predicted value : 2.331



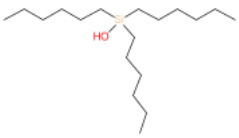
Compound #4

CAS: 947-04-6
Dataset id:219 (Training Set)
SMILES: O=C1NCCCCCCCCCCC1
Similarity: 0.706
Experimental value : 0.41
Predicted value : 1.594



Compound #5

CAS: 504-03-0
Dataset id:635 (Test Set)
SMILES: N1C(C)CCCC1C
Similarity: 0.687
Experimental value : 0.7
Predicted value : 1.002



Compound #6

CAS: 60782-58-3
Dataset id:518 (Training Set)
SMILES: O[Si](CCCCCC)(CCCCCC)CCCCCC
Similarity: 0.685
Experimental value : 2.7
Predicted value : 2.098

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.471 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.785 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✔	Accuracy of prediction for similar molecules Accuracy index = 0.359 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✖	Concordance for similar molecules Concordance index = 1.165 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 = 0.552 Explanation: the maximum error in prediction of similar molecules found in the training set has a moderate 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 = 0.6 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:

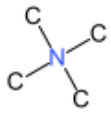
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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:

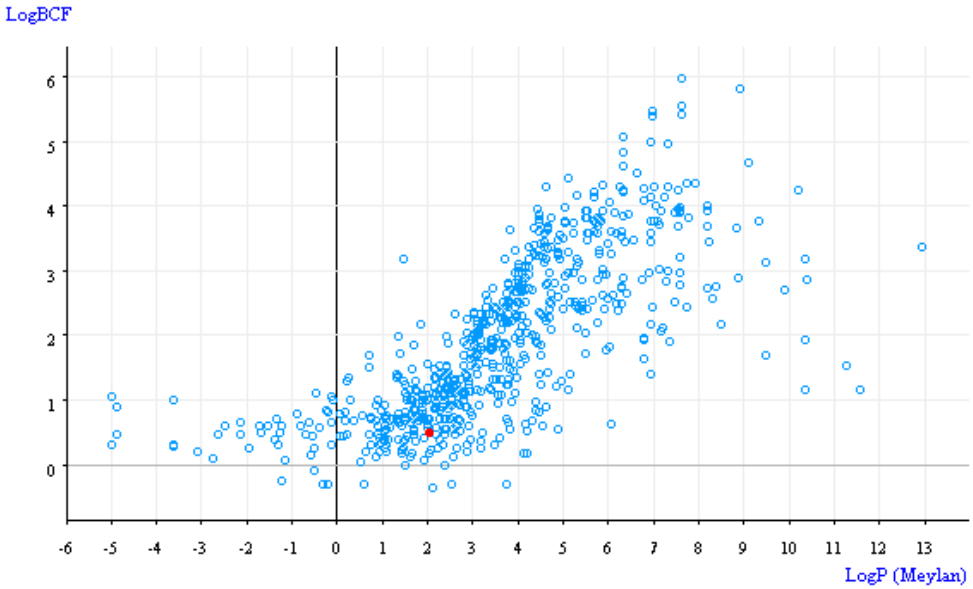


Fragment defined by the SMILES: CN(C)(C)C
The fragment has never been found in the model's training set

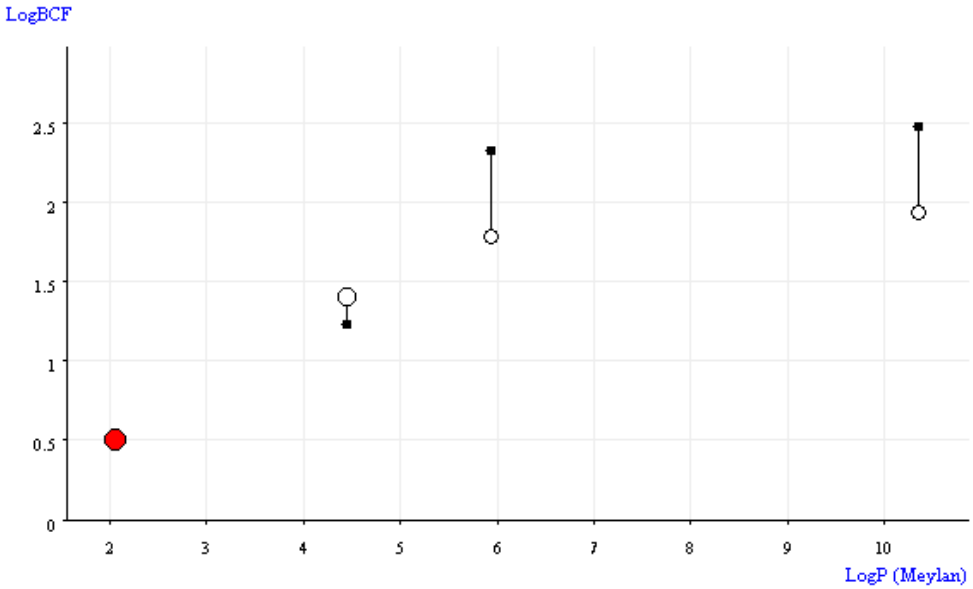
4.2 Reasoning:
Analysis of Molecular Descriptors

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.

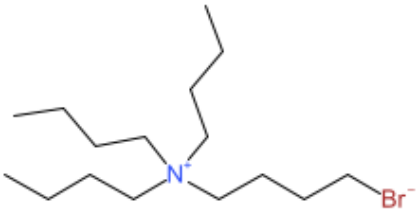


Following, a scatterplot of LogP (Meylan) against response values only for 3 most similar compounds in the training set. Red dot is the value of the studied compound, black outlined circles represents experimental values of compounds from training set, black dots represents predicted value of the same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



1. Prediction Summary

Prediction for compound Molecule 0 -

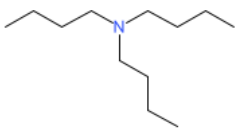


Prediction: Reliability:

Prediction is 1.07 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
- 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
- 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)

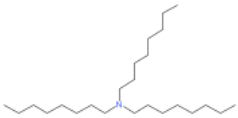
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted BCF (up) [log(L/kg)]: 1.07
Predicted BCF (up) [L/kg]: 12
Predicted BCF (low) [log(L/kg)]: 0.87
Predicted BCF (low) [L/kg]: 7
Predicted BCF (mid) [log(L/kg)]: 0.92
Predicted BCF (mid) [L/kg]: 8
Predicted LogP (Meylan/Kowwin): 2.05
Predicted LogP reliability: Low
Predicted kM (Meylan): -0.3
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



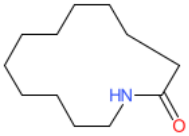
Compound #1

CAS: 102-82-9
Dataset id:781 (Training Set)
SMILES: N(CCCC)(CCCC)CCCC
Similarity: 0.819
Experimental value : 1.2
Predicted value : 2.232



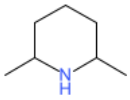
Compound #2

CAS: 1116-76-3
Dataset id:72 (Training Set)
SMILES: N(CCCCCCCC)(CCCCCCCC)CCCCCCCC
Similarity: 0.757
Experimental value : 1.86
Predicted value : 0.412



Compound #3

CAS: 947-04-6
Dataset id:52 (Training Set)
SMILES: O=C1NCCCCCCCCCCC1
Similarity: 0.706
Experimental value : 0.205
Predicted value : 1.845

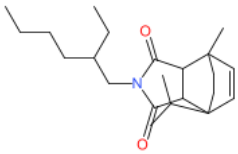


Compound #4

CAS: 504-03-0
Dataset id:384 (Training Set)
SMILES: N1C(C)CCCC1C
Similarity: 0.687
Experimental value : 0.2
Predicted value : 1.066

Compound #5

CAS: 140-72-7
Dataset id:579 (Training Set)
SMILES: c1cc[n+](cc1)CCCCCCCCCCCCCCCC
Similarity: 0.687
Experimental value : 1.543
Predicted value : 1.614



Compound #6

CAS: 13358-11-7
Dataset id:295 (Training Set)
SMILES: O=C1N(C(=O)C3C1C2(C=CC3(CC2)C(C)C)(C))CC(CC)CCCC
Similarity: 0.675
Experimental value : 3.055
Predicted value : 2.849

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.667 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.785 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✖	Accuracy of prediction for similar molecules Accuracy index = 1.24 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✔	Concordance for similar molecules Concordance index = 0.456 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 = 1.448 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 = 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:

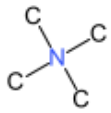
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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:

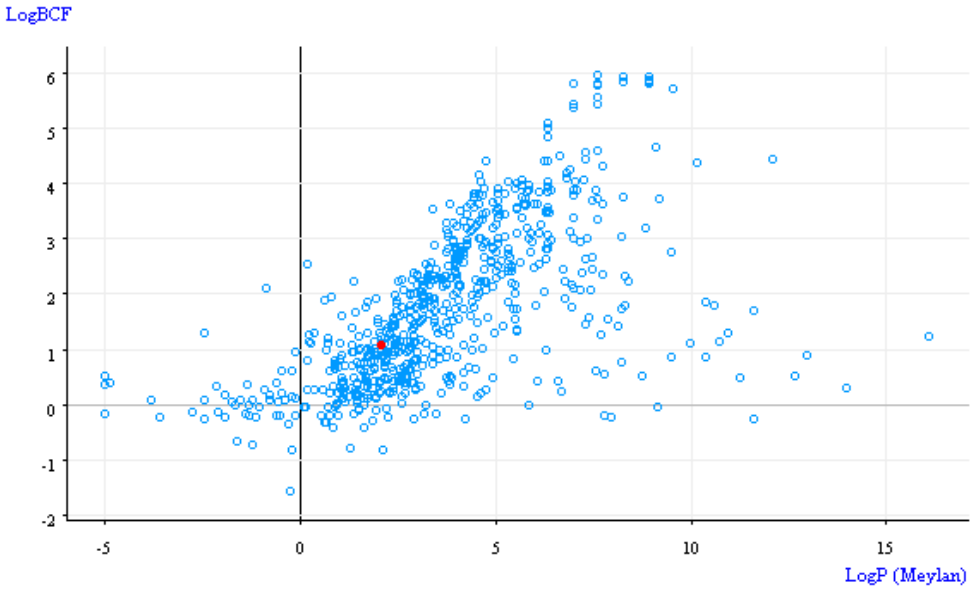


Fragment defined by the SMILES: CN(C)(C)C
The fragment has less than 3 occurrences in the model's training set

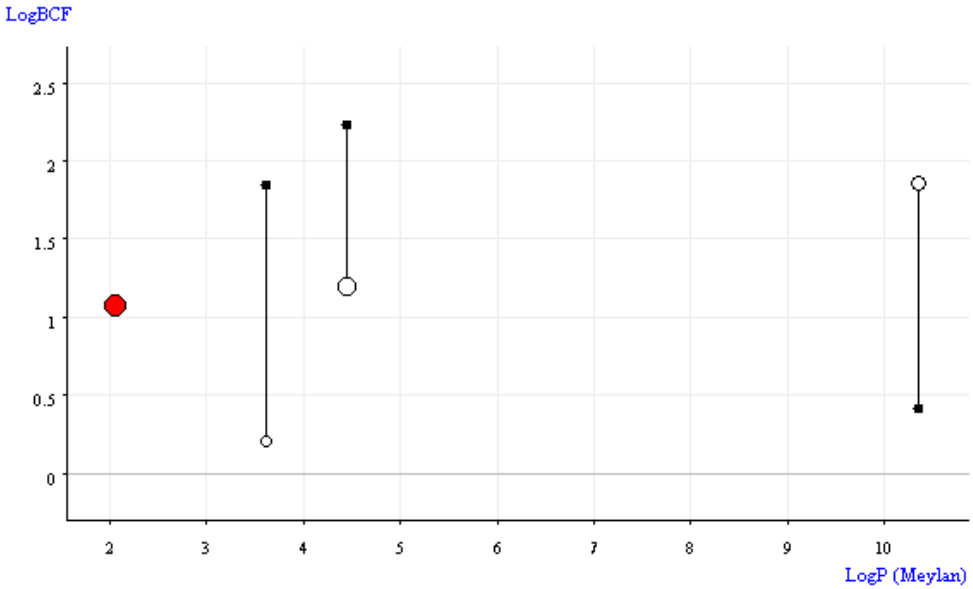
4.2 Reasoning:
Analysis of Molecular Descriptors

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.

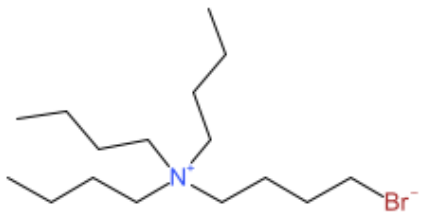


Following, a scatterplot of LogP (Meylan) against response values only for 3 most similar compounds in the training set. Red dot is the value of the studied compound, black outlined circles represents experimental values of compounds from training set, black dots represents predicted value of the same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



1. Prediction Summary

Prediction for compound Molecule 0 -

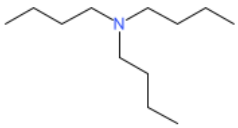


Prediction: Reliability:

Prediction is 2.02 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
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability
- 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)

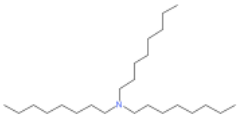
Compound: Molecule 0
Compound SMILES: CCCC[N+](CCCC)(CCCC)CCCC[Br-]
Experimental value: -
Predicted BCF [log(L/kg)]: 2.02
Molecules used for prediction: 4
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



Compound #1

CAS: 102-82-9
Dataset id:618 (Training Set)
SMILES: N(CCCC)(CCCC)CCCC
Similarity: 0.819
Experimental value : 1.199
Predicted value : 0.678

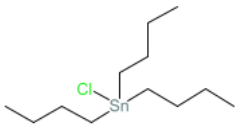


Compound #2

CAS: 1116-76-3
Dataset id:402 (Training Set)
SMILES: N(CCCCCCCC)(CCCCCCCC)CCCCCCCC
Similarity: 0.757
Experimental value : 1.969
Predicted value : 0.739

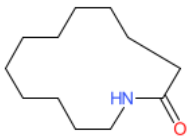
Compound #3

CAS: 4101-68-2
Dataset id:468 (Training Set)
SMILES: C(CCCCCBr)CCCCBr
Similarity: 0.756
Experimental value : 2.084
Predicted value : 3.14



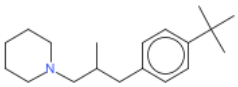
Compound #4

CAS: 1461-22-9
Dataset id:408 (Training Set)
SMILES: CCCC[Sn](CCCC)(CCCC)Cl
Similarity: 0.707
Experimental value : 3.296
Predicted value : 3.11



Compound #5

CAS: 947-04-6
Dataset id:396 (Training Set)
SMILES: O=C1NCCCCCCCCCCC1
Similarity: 0.706
Experimental value : 0.164
Predicted value : 0.807



Compound #6

CAS: 67306-00-7
Dataset id:656 (Training Set)
SMILES: c1cc(ccc1CC(C)CN2CCCCC2)C(C)(C)C
Similarity: 0.703
Experimental value : 2.204
Predicted value : 2.976

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.638 Explanation: The predicted compound is outside the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.75 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.748 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✓	Concordance for similar molecules Concordance index = 0.553 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 = 1.23 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 = 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:

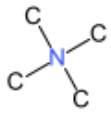
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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: CN(C)(C)C
The fragment has less than 3 occurrences in the model's training set

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