VEGA QSAR MODELS

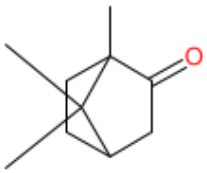
Report

Prediction and Applicability Domain analysis for models:

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

Core version: 1.3.18

Prediction for compound Molecule 0 -



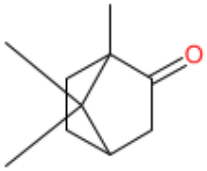
Prediction:

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

Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Used models: 4
Predicted Consensus Mutagen activity: NON-Mutagenic
Mutagenic Score: 0
Non-Mutagenic Score: 0.575
Model Caesar assessment: NON-Mutagenic (MODERATE reliability)
Model ISS assessment: NON-Mutagenic (LOW reliability)
Model SarPy assessment: NON-Mutagenic (MODERATE reliability)
Model KNN assessment: NON-Mutagenic (GOOD reliability)
Remarks:
none

1. Prediction Summary

Prediction for compound Molecule 0 -

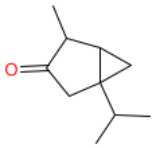


Prediction: Reliability:

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

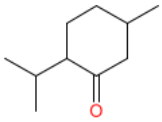
Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
Structural Alerts: -
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

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



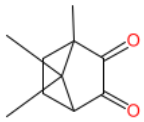
Compound #1

CAS: 546-80-5
Dataset id:4112 (Test Set)
SMILES: O=C1CC2(CC2(C1C))(C(C)C)
Similarity: 0.96
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



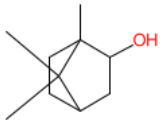
Compound #2

CAS: 89-80-5
Dataset id:303 (Training Set)
SMILES: O=C1CC(C)CCC1C(C)C
Similarity: 0.945
Experimental value : Mutagenic
Predicted value : NON-Mutagenic



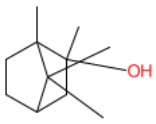
Compound #3

CAS: 465-29-2
Dataset id:3953 (Training Set)
SMILES: O=C1C(=O)C2(C)(CCC1C2(C)(C))
Similarity: 0.907
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



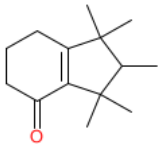
Compound #4

CAS: 464-43-7
Dataset id:1154 (Training Set)
SMILES: OC1CC2CCC1(C)C2(C)(C)
Similarity: 0.884
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #5

CAS: 2371-42-8
Dataset id:1595 (Training Set)
SMILES: OC1(C)(CC2CCC1(C)C2(C)(C))
Similarity: 0.868
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #6

CAS: 33704-61-9
Dataset id:2693 (Training Set)
SMILES: O=C1C2=C(CCC1)C(C)(C)C(C)C2(C)C
Similarity: 0.866
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

3.2 Applicability Domain:
Measured Applicability Domain Scores

⚠	Global AD Index AD index = 0.788 Explanation: The predicted compound could be out of the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.935 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.664 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 0.664 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 = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

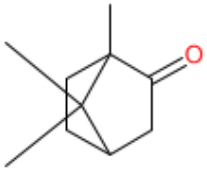
✓ The feature has a good assessment, model is reliable regarding this aspect.

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

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

1. Prediction Summary

Prediction for compound Molecule 0 -

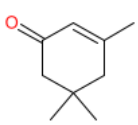


Prediction: Reliability:

Prediction is NON-Mutagenic, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Accuracy of prediction for similar molecules found in the training set is not adequate

Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
Structural Alerts: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

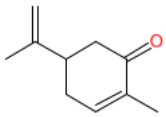
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 78-59-1
Dataset id:198 (Training Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848
Experimental value : NON-Mutagenic
Predicted value : Mutagenic

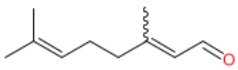
Alerts (not found also in the target): SA10 alfa, beta unsaturated carbonyls



Compound #2

CAS: 2244-16-8
Dataset id:662 (Training Set)
SMILES: O=C1C(=CCC(C(=C)C)C1)C
Similarity: 0.805
Experimental value : NON-Mutagenic
Predicted value : Mutagenic

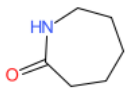
Alerts (not found also in the target): SA10 alfa, beta unsaturated carbonyls



Compound #3

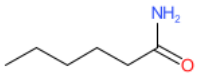
CAS: 5392-40-5
Dataset id:588 (Training Set)
SMILES: O=CC=C(C)CCC=C(C)C
Similarity: 0.732
Experimental value : NON-Mutagenic
Predicted value : Mutagenic

Alerts (not found also in the target): SA10 alfa, beta unsaturated carbonyls



Compound #4

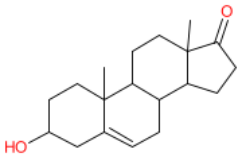
CAS: 105-60-2
Dataset id:78 (Training Set)
SMILES: O=C1NCCCCC1
Similarity: 0.729
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #5

CAS: 628-02-4
Dataset id:387 (Training Set)
SMILES: O=C(N)CCCCC
Similarity: 0.722
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

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



Compound #6

CAS: 53-43-0
Dataset id:836 (Training Set)
SMILES: O=C2CCC3C4CC=C1CC(O)CCC1(C)C4(CCC23(C))
Similarity: 0.719
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 = 0.825 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
✗	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 = 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:

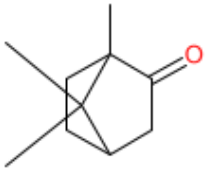
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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 -

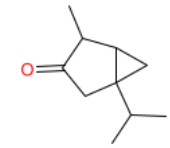


Prediction: Reliability:

Prediction is NON-Mutagenic, but the result shows some critical aspects, which require to be checked:
- Accuracy of prediction for similar molecules found in the training set is not optimal
- some similar molecules found in the training set have experimental values that disagree with the predicted value
The following relevant fragments have been found: SM162; SM163; SM169; SM177; SM182

Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
No. alerts for mutagenicity: 0
No. alerts for non-mutagenicity: 5
Structural Alerts: SM162; SM163; SM169; SM177; SM182
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

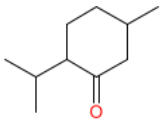
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 546-80-5
Dataset id:4112 (Test Set)
SMILES: O=C1CC2(CC2(C1C))(C(C)C)
Similarity: 0.96
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

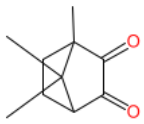
Alerts (found also in the target): SM163; SM169; SM177; SM182



Compound #2

CAS: 89-80-5
Dataset id:303 (Training Set)
SMILES: O=C1CC(C)CCC1C(C)C
Similarity: 0.945
Experimental value : Mutagenic
Predicted value : NON-Mutagenic

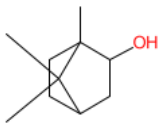
Alerts (found also in the target): SM163; SM169; SM177; SM182



Compound #3

CAS: 465-29-2
Dataset id:3953 (Training Set)
SMILES: O=C1C(=O)C2(C)(CCC1C2(C)(C))
Similarity: 0.907
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM162; SM163; SM169; SM177; SM182

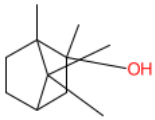


Compound #4

CAS: 464-43-7
Dataset id:1154 (Training Set)
SMILES: OC1CC2CCC1(C)C2(C)(C)
Similarity: 0.884
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM162; SM163; SM169; SM182

Alerts (not found also in the target): SM145



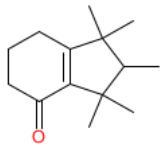
Compound #5

CAS: 2371-42-8
Dataset id:1595 (Training Set)
SMILES: OC1(C)(CC2CCC1(C)C2(C)(C))
Similarity: 0.868
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM162; SM163; SM169; SM182

Alerts (not found also in the target): SM145

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



Compound #6

CAS: 33704-61-9
Dataset id:2693 (Training Set)
SMILES: O=C1C2=C(CCC1)C(C)(C)C(C)C2(C)C
Similarity: 0.866
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM163; SM169; SM177; SM182

Alerts (not found also in the target): SM180

3.2 Applicability Domain:
Measured Applicability Domain Scores

⚠	Global AD Index AD index = 0.788 Explanation: The predicted compound could be out of the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.935 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.664 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 0.664 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.

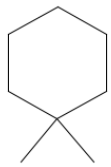
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

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

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

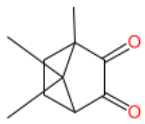
(Molecule 0) Reasoning on fragments/structural alerts - 1 of 5:.

Fragment found: SM162



Sarpy alert n. 162 for NON-Mutagenicity, defined by SMARTS: C1CC(CC(C1))(C)C

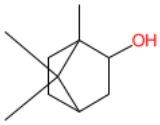
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 465-29-2
Dataset id:3953 (Training Set)
SMILES: O=C1C(=O)C2(C)(CCC1C2(C)(C))
Similarity: 0.907

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

Alerts (found also in the target): SM162; SM163; SM169; SM177; SM182

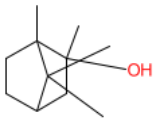


CAS: 464-43-7
Dataset id:1154 (Training Set)
SMILES: OC1CC2CCC1(C)C2(C)(C)
Similarity: 0.884

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

Alerts (found also in the target): SM162; SM163; SM169; SM182

Alerts (not found also in the target): SM145



CAS: 2371-42-8
Dataset id:1595 (Training Set)
SMILES: OC1(C)(CC2CCC1(C)C2(C)(C))
Similarity: 0.868

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

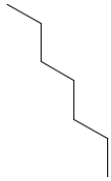
Alerts (found also in the target): SM162; SM163; SM169; SM182

Alerts (not found also in the target): SM145

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

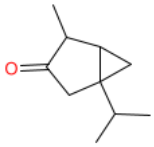
(Molecule 0) Reasoning on fragments/structural alerts - 2 of 5:.

Fragment found: SM163



Sarpy alert n. 163 for NON-Mutagenicity, defined by SMARTS: CCCCCCC

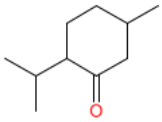
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 546-80-5
Dataset id:4112 (Test Set)
SMILES: O=C1CC2(CC2(C1C))(C(C)C)
Similarity: 0.96

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

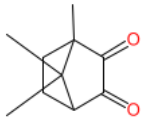
Alerts (found also in the target): SM163; SM169; SM177; SM182



CAS: 89-80-5
Dataset id:303 (Training Set)
SMILES: O=C1CC(C)CCC1C(C)C
Similarity: 0.945

Experimental value : Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM163; SM169; SM177; SM182



CAS: 465-29-2
Dataset id:3953 (Training Set)
SMILES: O=C1C(=O)C2(C)(CCC1C2(C)(C))
Similarity: 0.907

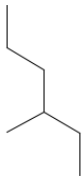
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM162; SM163; SM169; SM177; SM182

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

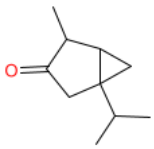
(Molecule 0) Reasoning on fragments/structural alerts - 3 of 5:.

Fragment found: SM169



Sarpy alert n. 169 for NON-Mutagenicity, defined by SMARTS: CC(CC)CCC

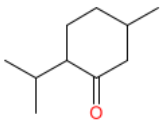
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 546-80-5
Dataset id:4112 (Test Set)
SMILES: O=C1CC2(CC2(C1C))(C(C)C)
Similarity: 0.96

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

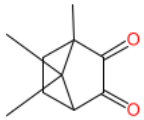
Alerts (found also in the target): SM163; SM169; SM177; SM182



CAS: 89-80-5
Dataset id:303 (Training Set)
SMILES: O=C1CC(C)CCC1C(C)C
Similarity: 0.945

Experimental value : Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM163; SM169; SM177; SM182



CAS: 465-29-2
Dataset id:3953 (Training Set)
SMILES: O=C1C(=O)C2(C)(CCC1C2(C)(C))
Similarity: 0.907

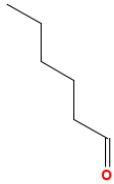
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM162; SM163; SM169; SM177; SM182

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

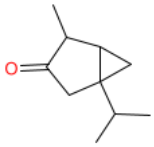
(Molecule 0) Reasoning on fragments/structural alerts - 4 of 5:.

Fragment found: SM177



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

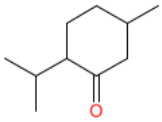
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 546-80-5
Dataset id:4112 (Test Set)
SMILES: O=C1CC2(CC2(C1C))(C(C)C)
Similarity: 0.96

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

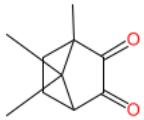
Alerts (found also in the target): SM163; SM169; SM177; SM182



CAS: 89-80-5
Dataset id:303 (Training Set)
SMILES: O=C1CC(C)CCC1C(C)C
Similarity: 0.945

Experimental value : Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM163; SM169; SM177; SM182



CAS: 465-29-2
Dataset id:3953 (Training Set)
SMILES: O=C1C(=O)C2(C)(CCC1C2(C)(C))
Similarity: 0.907

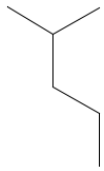
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM162; SM163; SM169; SM177; SM182

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

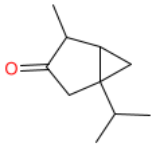
(Molecule 0) Reasoning on fragments/structural alerts - 5 of 5:.

Fragment found: SM182



Sarpy alert n. 182 for NON-Mutagenicity, defined by SMARTS: CCCC(C)C

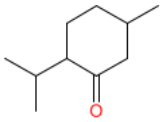
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 546-80-5
Dataset id:4112 (Test Set)
SMILES: O=C1CC2(CC2(C1C))(C(C)C)
Similarity: 0.96

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

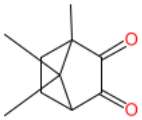
Alerts (found also in the target): SM163; SM169; SM177; SM182



CAS: 89-80-5
Dataset id:303 (Training Set)
SMILES: O=C1CC(C)CCC1C(C)C
Similarity: 0.945

Experimental value : Mutagenic
Predicted value : NON-Mutagenic

Alerts (found also in the target): SM163; SM169; SM177; SM182



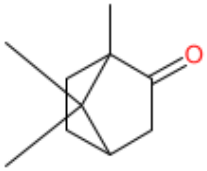
CAS: 465-29-2
Dataset id:3953 (Training Set)
SMILES: O=C1C(=O)C2(C)(CCC1C2(C)(C))
Similarity: 0.907

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

Alerts (found also in the target): SM162; SM163; SM169; SM177; SM182

1. Prediction Summary

Prediction for compound Molecule 0 -

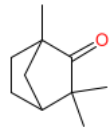


Prediction: Reliability:

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

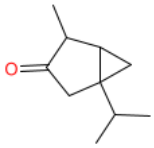
Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Mutagen activity: NON-Mutagenic
Molecules used for prediction: 4
Reliability: The predicted compound is into the Applicability Domain of the model
Remarks:
none

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



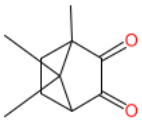
Compound #1

CAS: 1195-79-5
Dataset id:801 (Training Set)
SMILES: O=C2C1(C)(CCC(C1)C2(C)C)
Similarity: 0.995
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



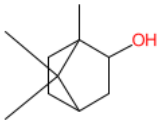
Compound #2

CAS: 546-80-5
Dataset id:3567 (Training Set)
SMILES: O=C1CC2(CC2(C1C))(C(C)C)
Similarity: 0.96
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



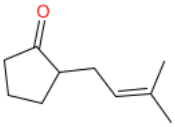
Compound #3

CAS: 10373-78-1
Dataset id:149 (Training Set)
SMILES: O=C1C(=O)C2(C)(CCC1C2(C)(C))
Similarity: 0.907
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



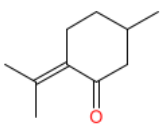
Compound #4

CAS: 507-70-0
Dataset id:3316 (Training Set)
SMILES: OC1CC2CCC1(C)C2(C)(C)
Similarity: 0.884
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #5

CAS: 2520-60-7
Dataset id:2430 (Training Set)
SMILES: O=C1CCCC1CC=C(C)C
Similarity: 0.878
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic



Compound #6

CAS: 89-82-7
Dataset id:5124 (Training Set)
SMILES: O=C1C(=C(C)C)CCC(C)C1
Similarity: 0.874
Experimental value : NON-Mutagenic
Predicted value : NON-Mutagenic

3.2 Applicability Domain:
Measured Applicability Domain Scores

✓	Global AD Index AD index = 0.962 Explanation: The predicted compound is into the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.925 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:

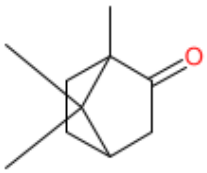
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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 -

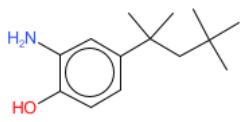


Prediction: Reliability:

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
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)

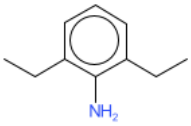
Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Mutagen activity: NA
Structural Alerts: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



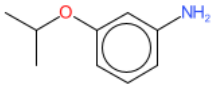
Compound #1

CAS: N.A.
Dataset id:4177 (Training Set)
SMILES: Oc1ccc(cc1(N))C(C)(C)CC(C)(C)C
Similarity: 0.62
Experimental value : NON-Mutagenic
Predicted value : Mutagenic



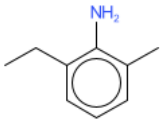
Compound #2

CAS: N.A.
Dataset id:2394 (Training Set)
SMILES: Nc1c(cccc1CC)CC
Similarity: 0.602
Experimental value : Mutagenic
Predicted value : Mutagenic



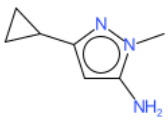
Compound #3

CAS: N.A.
Dataset id:3753 (Training Set)
SMILES: O(c1cccc(N)c1)C(C)C
Similarity: 0.597
Experimental value : NON-Mutagenic
Predicted value : Mutagenic



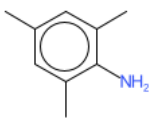
Compound #4

CAS: N.A.
Dataset id:2759 (Training Set)
SMILES: Nc1c(cccc1CC)C
Similarity: 0.578
Experimental value : Mutagenic
Predicted value : Mutagenic



Compound #5

CAS: N.A.
Dataset id:3520 (Training Set)
SMILES: n1c(cc(N)n1C)C2CC2
Similarity: 0.578
Experimental value : Mutagenic
Predicted value : NA



Compound #6

CAS: N.A.
Dataset id:1077 (Training Set)
SMILES: Nc1c(cc(cc1C)C)C
Similarity: 0.577
Experimental value : Mutagenic
Predicted value : Mutagenic

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.611 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✖	Accuracy of prediction for similar molecules Accuracy index = 0.485 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 = 0.85 Explanation: some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)..

Symbols explanation:

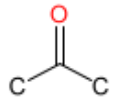
✔ The feature has a good assessment, model is reliable regarding this aspect.

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

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

4.1 Reasoning:
Relevant Chemical Fragments and Moieties

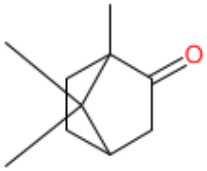
(Molecule 0) Reasoning on rare and missing Atom Centered Fragments .
The following Atom Centered Fragments have been found in the molecule, but they are not found or rarely found in the model's training set:



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

1. Prediction Summary

Prediction for compound Molecule 0 -

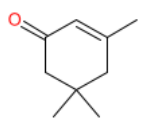


Prediction: Reliability:

Prediction is NON-Carcinogen, 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
- similar molecules found in the training set have experimental values that disagree with the predicted value

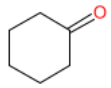
Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Carcinogen activity: NON-Carcinogen
P(Carcinogen): 0.13
P(NON-Carcinogen): 0.87
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

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



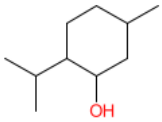
Compound #1

CAS: 78-59-1
Dataset id:404 (Training Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848
Experimental value : Carcinogen
Predicted value : Carcinogen



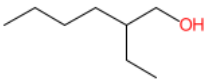
Compound #2

CAS: 108-94-1
Dataset id:187 (Test Set)
SMILES: O=C1CCCCC1
Similarity: 0.835
Experimental value : NON-Carcinogen
Predicted value : Carcinogen



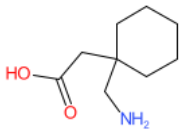
Compound #3

CAS: 89-78-1
Dataset id:427 (Training Set)
SMILES: OC1CC(C)CCC1(C(C)C)
Similarity: 0.827
Experimental value : NON-Carcinogen
Predicted value : Carcinogen



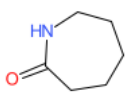
Compound #4

CAS: 104-76-7
Dataset id:314 (Training Set)
SMILES: OCC(CC)CCCC
Similarity: 0.784
Experimental value : NON-Carcinogen
Predicted value : Carcinogen



Compound #5

CAS: 60142-96-3
Dataset id:48 (Training Set)
SMILES: O=C(O)CC1(CN)(CCCCC1)
Similarity: 0.763
Experimental value : Carcinogen
Predicted value : Carcinogen



Compound #6

CAS: 105-60-2
Dataset id:124 (Training Set)
SMILES: O=C1NCCCCC1
Similarity: 0.729
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

⚠	Global AD Index AD index = 0.649 Explanation: The predicted compound could be out of the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.841 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.505 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✗	Concordance for similar molecules Concordance index = 0.495 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..
✓	Model class assignment reliability Pos/Non-Pos difference = 0.739 Explanation: model class assignment is well defined..
✓	Neural map neurons concordance Neurons concordance = 1 Explanation: predicted value agrees with experimental values of training set compounds laying in the same neuron..

Symbols explanation:

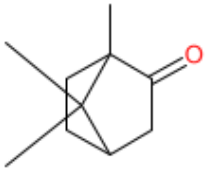
✓ The feature has a good assessment, model is reliable regarding this aspect.

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

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

1. Prediction Summary

Prediction for compound Molecule 0 -

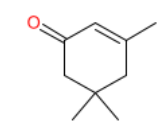


Prediction: Reliability:

Prediction is NON-Carcinogen, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Accuracy of prediction for similar molecules found in the training set is not optimal
- similar molecules found in the training set have experimental values that disagree with the predicted value

Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Carcinogen activity: NON-Carcinogen
Structural Alerts: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

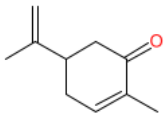
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 78-59-1
Dataset id:198 (Training Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848
Experimental value : Carcinogen
Predicted value : Carcinogen

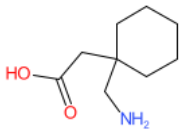
Alerts (not found also in the target): SA10 alfa, beta unsaturated carbonyls



Compound #2

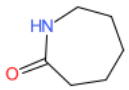
CAS: 2244-16-8
Dataset id:662 (Training Set)
SMILES: O=C1C(=CCC(C(=C)C)C1)C
Similarity: 0.805
Experimental value : NON-Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): SA10 alfa, beta unsaturated carbonyls



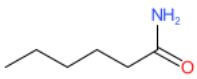
Compound #3

CAS: 60142-96-3
Dataset id:789 (Training Set)
SMILES: O=C(O)CC1(CN)(CCCCC1)
Similarity: 0.763
Experimental value : Carcinogen
Predicted value : NON-Carcinogen



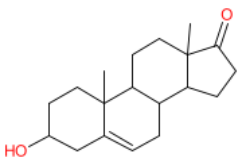
Compound #4

CAS: 105-60-2
Dataset id:78 (Training Set)
SMILES: O=C1NCCCCC1
Similarity: 0.729
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #5

CAS: 628-02-4
Dataset id:387 (Training Set)
SMILES: O=C(N)CCCCC
Similarity: 0.722
Experimental value : Carcinogen
Predicted value : NON-Carcinogen



Compound #6

CAS: 53-43-0
Dataset id:836 (Training Set)
SMILES: O=C2CCC3C4CC=C1CC(O)CCC1(C)C4(CCC23(C))
Similarity: 0.719
Experimental value : Carcinogen
Predicted value : NON-Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.642 Explanation: The predicted compound is outside the Applicability Domain of the model.
✔	Similar molecules with known experimental value Similarity index = 0.825 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.516 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✖	Concordance for similar molecules Concordance index = 0.484 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:

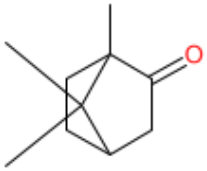
✔ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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.

1. Prediction Summary

Prediction for compound Molecule 0 -

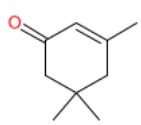


Prediction: Reliability:

Prediction is Possible NON-Carcinogen, the result appears reliable.
Anyhow, you should check it through the evaluation of the information given in the following sections.Anyway some issues could be not optimal:
- some similar molecules found in the training set have experimental values that disagree with the predicted value

Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Carcinogenic activity: Possible NON-Carcinogen
No. alerts for carcinogenicity: 0
Structural Alerts: -
Reliability: The predicted compound is into the Applicability Domain of the model
Remarks:
none

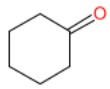
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

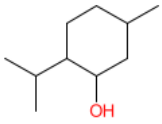
CAS: 78-59-1
Dataset id:159 (Training Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848
Experimental value : Carcinogen
Predicted value : Carcinogen

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



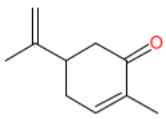
Compound #2

CAS: 108-94-1
Dataset id:934 (Training Set)
SMILES: O=C1CCCCC1
Similarity: 0.835
Experimental value : NON-Carcinogen
Predicted value : Possible NON-Carcinogen



Compound #3

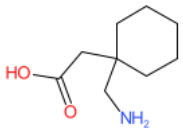
CAS: 15356-70-4
Dataset id:654 (Training Set)
SMILES: OC1CC(C)CCC1(C(C)C)
Similarity: 0.827
Experimental value : NON-Carcinogen
Predicted value : Possible NON-Carcinogen



Compound #4

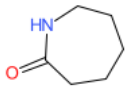
CAS: 2244-16-8
Dataset id:690 (Training Set)
SMILES: O=C1C(=CCC(C(=C)C)C1)C
Similarity: 0.805
Experimental value : NON-Carcinogen
Predicted value : Carcinogen

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



Compound #5

CAS: 60142-96-3
Dataset id:725 (Training Set)
SMILES: O=C(O)CC1(CN)(CCCCC1)
Similarity: 0.763
Experimental value : Carcinogen
Predicted value : Possible NON-Carcinogen



Compound #6

CAS: 105-60-2
Dataset id:63 (Training Set)
SMILES: O=C1NCCCCC1
Similarity: 0.729
Experimental value : NON-Carcinogen
Predicted value : Possible NON-Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✓	Global AD Index AD index = 0.825 Explanation: The predicted compound is into the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.836 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 = 0.661 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

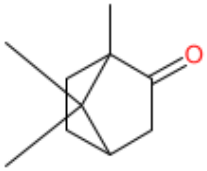
✓ The feature has a good assessment, model is reliable regarding this aspect.

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

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

1. Prediction Summary

Prediction for compound Molecule 0 -

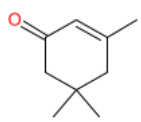


Prediction: Reliability:

Prediction is Possible NON-Carcinogen, the result appears reliable.
Anyhow, you should check it through the evaluation of the information given in the following sections.Anyway some issues could be not optimal:
- some similar molecules found in the training set have experimental values that disagree with the predicted value

Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Carcinogenic activity: Possible NON-Carcinogen
No. alerts for carcinogenicity: 0
Structural Alerts: -
Reliability: The predicted compound is into the Applicability Domain of the model
Remarks:
none

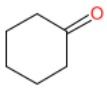
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

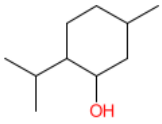
CAS: 78-59-1
Dataset id:404 (Training Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848
Experimental value : Carcinogen
Predicted value : Carcinogen

Alerts (not found also in the target): Carcinogenity alert no. 85



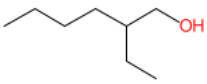
Compound #2

CAS: 108-94-1
Dataset id:187 (Test Set)
SMILES: O=C1CCCCC1
Similarity: 0.835
Experimental value : NON-Carcinogen
Predicted value : Possible NON-Carcinogen



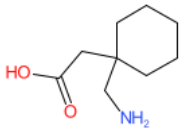
Compound #3

CAS: 89-78-1
Dataset id:427 (Training Set)
SMILES: OC1CC(C)CCC1(C(C)C)
Similarity: 0.827
Experimental value : NON-Carcinogen
Predicted value : Possible NON-Carcinogen



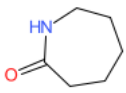
Compound #4

CAS: 104-76-7
Dataset id:314 (Training Set)
SMILES: OCC(CC)CCCC
Similarity: 0.784
Experimental value : NON-Carcinogen
Predicted value : Possible NON-Carcinogen



Compound #5

CAS: 60142-96-3
Dataset id:48 (Training Set)
SMILES: O=C(O)CC1(CN)(CCCCC1)
Similarity: 0.763
Experimental value : Carcinogen
Predicted value : Possible NON-Carcinogen



Compound #6

CAS: 105-60-2
Dataset id:124 (Training Set)
SMILES: O=C1NCCCCC1
Similarity: 0.729
Experimental value : NON-Carcinogen
Predicted value : Possible NON-Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✓	Global AD Index AD index = 0.825 Explanation: The predicted compound is into the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.836 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 = 0.661 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

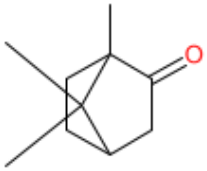
✓ The feature has a good assessment, model is reliable regarding this aspect.

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

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

1. Prediction Summary

Prediction for compound Molecule 0 -

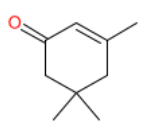


Prediction: Reliability:

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

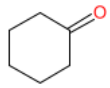
Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Oral Carcinogenic class: NON-Carcinogen
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

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



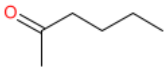
Compound #1

CAS: 78-59-1
Dataset id:185 (Training Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848
Experimental value : Carcinogen
Predicted value : NON-Carcinogen



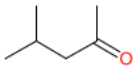
Compound #2

CAS: 108-94-1
Dataset id:416 (Training Set)
SMILES: O=C1CCCCC1
Similarity: 0.835
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



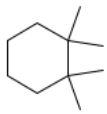
Compound #3

CAS: 591-78-6
Dataset id:542 (Training Set)
SMILES: O=C(C)CCCC
Similarity: 0.823
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



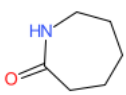
Compound #4

CAS: 108-10-1
Dataset id:581 (Test Set)
SMILES: O=C(C)CC(C)C
Similarity: 0.807
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #5

CAS: 30501-43-0
Dataset id:692 (Training Set)
SMILES: CC1(C)(CCCCC1(C)(C))
Similarity: 0.78
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #6

CAS: 105-60-2
Dataset id:369 (Training Set)
SMILES: O=C1NCCCCC1
Similarity: 0.729
Experimental value : NON-Carcinogen
Predicted value : Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

⚠	Global AD Index AD index = 0.646 Explanation: The predicted compound could be out of the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.841 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
✗	Accuracy of prediction for similar molecules Accuracy index = 0.495 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✗	Concordance for similar molecules Concordance index = 0.495 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:

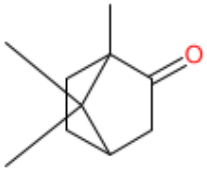
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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.

1. Prediction Summary

Prediction for compound Molecule 0 -

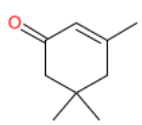


Prediction: Reliability:

Prediction is -0.97, 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

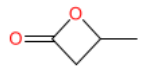
Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Oral Carcinogenicity SF (log form) [log(1/(mg/kg-day))]: -0.97
Predicted Oral Carcinogenicity SF [1/(mg/kg-day)]: 0.1073
Experimental value [1/(mg/kg-day)]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



Compound #1

CAS: 78-59-1
Dataset id:185 (Test Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848
Experimental value : -3.02
Predicted value : -1.487

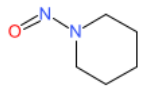


Compound #2

CAS: 3068-88-0
Dataset id:53 (Training Set)
SMILES: O=C1OC(C)C1
Similarity: 0.684
Experimental value : 0
Predicted value : -0.807

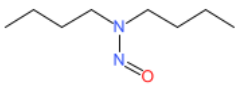
Compound #3

CAS: 1634-04-4
Dataset id:192 (Training Set)
SMILES: O(C)C(C)(C)C
Similarity: 0.683
Experimental value : -2.74
Predicted value : -1.149



Compound #4

CAS: 100-75-4
Dataset id:234 (Training Set)
SMILES: O=NN1CCCCC1
Similarity: 0.66
Experimental value : 0.97
Predicted value : 1.097



Compound #5

CAS: 924-16-3
Dataset id:224 (Test Set)
SMILES: O=NN(CCCC)CCCC
Similarity: 0.652
Experimental value : 0.73
Predicted value : 0.473

Compound #6

CAS: 75-56-9
Dataset id:268 (Training Set)
SMILES: O1CC1C
Similarity: 0.648
Experimental value : -0.62
Predicted value : -0.91

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.7 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.745 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 1.17 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✖	Concordance for similar molecules Concordance index = 1.51 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 = 1.533 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✔	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✔	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

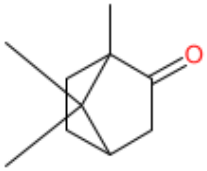
✔ The feature has a good assessment, model is reliable regarding this aspect.

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

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

1. Prediction Summary

Prediction for compound Molecule 0 -

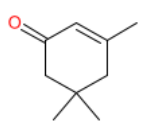


Prediction: Reliability:

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

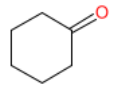
Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Inhalation Carcinogenic class: NON-Carcinogen
Reliability: The predicted compound is into the Applicability Domain of the model
Remarks:
none

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



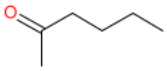
Compound #1

CAS: 78-59-1
Dataset id:529 (Training Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



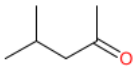
Compound #2

CAS: 108-94-1
Dataset id:379 (Training Set)
SMILES: O=C1CCCCC1
Similarity: 0.835
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



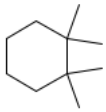
Compound #3

CAS: 591-78-6
Dataset id:518 (Training Set)
SMILES: O=C(C)CCCC
Similarity: 0.823
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



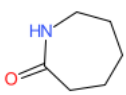
Compound #4

CAS: 108-10-1
Dataset id:560 (Test Set)
SMILES: O=C(C)CC(C)C
Similarity: 0.807
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #5

CAS: 30501-43-0
Dataset id:687 (Training Set)
SMILES: CC1(C)(CCCCC1(C)(C))
Similarity: 0.78
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen



Compound #6

CAS: 105-60-2
Dataset id:324 (Training Set)
SMILES: O=C1NCCCCC1
Similarity: 0.729
Experimental value : NON-Carcinogen
Predicted value : NON-Carcinogen

3.2 Applicability Domain:
Measured Applicability Domain Scores

✓	Global AD Index AD index = 0.917 Explanation: The predicted compound is into the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.841 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
✓	Accuracy of prediction for similar molecules Accuracy index = 1 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
✓	Concordance for similar molecules Concordance index = 1 Explanation: Similar molecules found in the training set have experimental values that agree with the predicted value..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

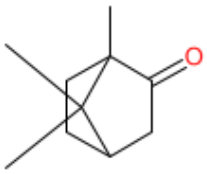
✓ The feature has a good assessment, model is reliable regarding this aspect.

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

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

1. Prediction Summary

Prediction for compound Molecule 0 -

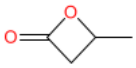


Prediction: Reliability:

Prediction is 0.31, but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- No similar compounds with known experimental value in the training set have been found
- Accuracy of prediction for similar molecules found in the training set is not optimal
- similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability
- some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments (1 infrequent_fragments found)

Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted Inhalation Carcinogenicity SF (log form) [log(1/(mg/kg-day))]: 0.31
Predicted Inhalation Carcinogenicity SF [1/(mg/kg-day)]: 2.05
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

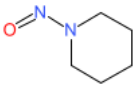


Compound #1

CAS: 3068-88-0
Dataset id:45 (Training Set)
SMILES: O=C1OC(C)C1
Similarity: 0.684
Experimental value : 0.01
Predicted value : 0.684

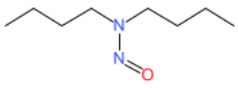
Compound #2

CAS: 1634-04-4
Dataset id:161 (Training Set)
SMILES: O(C)C(C)(C)C
Similarity: 0.683
Experimental value : -3.04
Predicted value : -0.844



Compound #3

CAS: 100-75-4
Dataset id:202 (Training Set)
SMILES: O=NN1CCCCC1
Similarity: 0.66
Experimental value : 0.98
Predicted value : 0.053

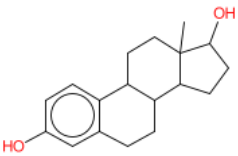


Compound #4

CAS: 924-16-3
Dataset id:192 (Training Set)
SMILES: O=NN(CCCC)CCCC
Similarity: 0.652
Experimental value : 0.75
Predicted value : -0.335

Compound #5

CAS: 75-56-9
Dataset id:228 (Training Set)
SMILES: O1CC1C
Similarity: 0.648
Experimental value : -1.89
Predicted value : -0.31



Compound #6

CAS: 50-28-2
Dataset id:116 (Training Set)
SMILES: Oc1ccc2c(c1)CCC3C2CCC4(C)(C(O)CCC34)
Similarity: 0.636
Experimental value : 1.59
Predicted value : 2.033

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.581 Explanation: The predicted compound is outside the Applicability Domain of the model.
✖	Similar molecules with known experimental value Similarity index = 0.683 Explanation: No similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 1.435 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✖	Concordance for similar molecules Concordance index = 1.826 Explanation: similar molecules found in the training set have experimental values that disagree with the predicted value..
✖	Maximum error of prediction among similar molecules Max error index = 2.196 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.

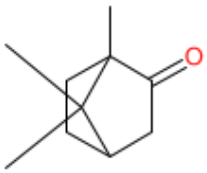
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

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

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

1. Prediction Summary

Prediction for compound Molecule 0 -

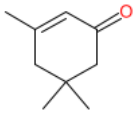


Prediction: Reliability:

Prediction is -3.0428, 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

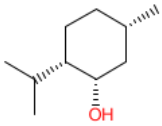
Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted TD50 [-log(mg/kg bw/day)]: -3.0428
Predicted TD50 [mg/kg bw/day]: 1103.45
Experimental TD50 [mg/kg bw/day]: -
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

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



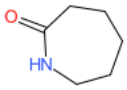
Compound #1

CAS: N.A.
Dataset id:146 (Training Set)
SMILES: C1(=CC(=O)CC(C1)(C)C)C
Similarity: 0.848
Experimental value : -3.958
Predicted value : -4.018



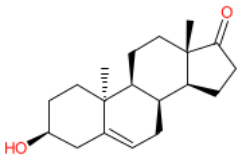
Compound #2

CAS: N.A.
Dataset id:200 (Test Set)
SMILES: C1[C@@H](C[C@@H]([C@@H](C1)C(C)C)O)C
Similarity: 0.827
Experimental value : -3.755
Predicted value : -3.627



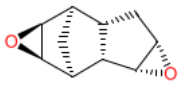
Compound #3

CAS: N.A.
Dataset id:141 (Training Set)
SMILES: C1CCNC(=O)CC1
Similarity: 0.729
Experimental value : -3.755
Predicted value : -2.929



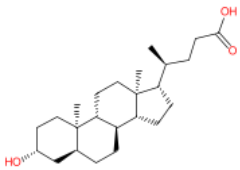
Compound #4

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: 0.719
Experimental value : -1.833
Predicted value : -0.788



Compound #5

CAS: N.A.
Dataset id:137 (Training Set)
SMILES:
C1[C@@H]2[C@@H]([C@H]3[C@@H]4[C@@H](O4)[C@@H]2C3)[C@H]2O[C@@H]12
Similarity: 0.695
Experimental value : -3.592
Predicted value : -2.37



Compound #6

CAS: N.A.
Dataset id:129 (Training Set)
SMILES:
[C@@H]1(C[C@H]2[C@@](CC1)([C@H]1[C@H](CC2)[C@@H]2[C@@](CC1)([C@@H](CC2)[C@@H](C)CCC(=O)O)C)C)O
Similarity: 0.69
Experimental value : -3.352
Predicted value : -0.763

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 = 0.837 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✓	Accuracy of prediction for similar molecules Accuracy index = 0.094 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
⚠	Concordance for similar molecules Concordance index = 0.814 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✓	Maximum error of prediction among similar molecules Max error index = 0.128 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 = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

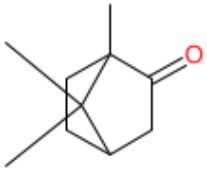
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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 -

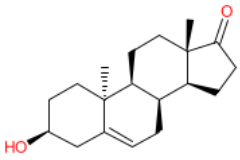


Prediction: Reliability:

Prediction is -2.9265, 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

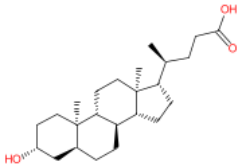
Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted TD50 [-log(mg/kg bw/day)]: -2.9265
Predicted TD50 [mg/kg bw/day]: 844.35
Experimental TD50 [mg/kg bw/day]: -
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

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



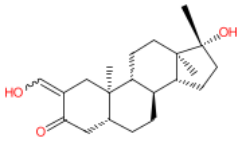
Compound #1

CAS: N.A.
Dataset id: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: 0.719
Experimental value : -1.921
Predicted value : -4.971



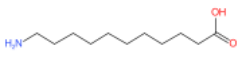
Compound #2

CAS: N.A.
Dataset id:47 (Training Set)
SMILES:
[C@@H]1(C[C@H]2[C@@](CC1)([C@H]1[C@H](CC2)[C@@H]2[C@@](CC1)([C@@H](CC2)[C@@H](C)CCC(=O)O)C)C)O
Similarity: 0.69
Experimental value : -3.276
Predicted value : -4.644



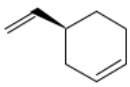
Compound #3

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: 0.687
Experimental value : -2.279
Predicted value : -3.484



Compound #4

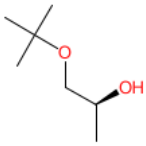
CAS: N.A.
Dataset id:117 (Training Set)
SMILES: C(CCCCCCCCCC(=O)O)N
Similarity: 0.673
Experimental value : -4.649
Predicted value : -4.782



Compound #5

CAS: N.A.
Dataset id:77 (Training Set)
SMILES: C1[C@H](CCC=C1)C=C
Similarity: 0.671
Experimental value : -5.121
Predicted value : -3.477

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



Compound #6

CAS: N.A.
Dataset id:112 (Training Set)
SMILES: C[C@@H](COC(C)(C)C)O
Similarity: 0.665
Experimental value : -4.021
Predicted value : -2.558

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.7 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.704 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✗	Accuracy of prediction for similar molecules Accuracy index = 2.209 Explanation: Accuracy of prediction for similar molecules found in the training set is not adequate..
✓	Concordance for similar molecules Concordance index = 0.678 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 = 3.05 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✓	Model's descriptors range check Descriptors range check = True Explanation: descriptors for this compound have values inside the descriptor range of the compounds of the training set..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

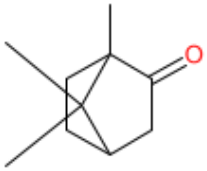
✓ The feature has a good assessment, model is reliable regarding this aspect.

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

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

1. Prediction Summary

Prediction for compound Molecule 0 -

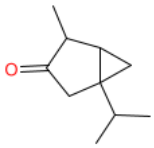


Prediction: Reliability:

Prediction is 1262.1 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

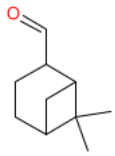
Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted log LD50 [log(mmol/Kg)]: 0.918
Predicted log LD50 [mg/Kg]: 1262.1
Molecules used for prediction: 3
Experimental value [mg/Kg]: -
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

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



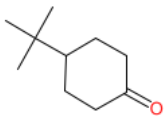
Compound #1

CAS: N.A.
Dataset id:233 (Training Set)
SMILES: O=C1CC2(CC2(C1C))(C(C)C)
Similarity: 0.96
Experimental value : 0.1
Predicted value : 1.039



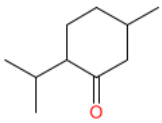
Compound #2

CAS: N.A.
Dataset id:5078 (Training Set)
SMILES: O=CC1CCC2CC1C2(C)(C)
Similarity: 0.951
Experimental value : 1.18
Predicted value : 0.675



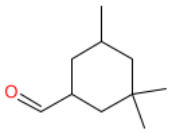
Compound #3

CAS: N.A.
Dataset id:264 (Training Set)
SMILES: O=C1CCC(CC1)C(C)(C)C
Similarity: 0.945
Experimental value : 1.51
Predicted value : 1.057



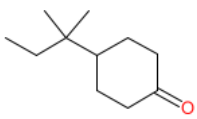
Compound #4

CAS: N.A.
Dataset id:5089 (Training Set)
SMILES: O=C1CC(C)CCC1C(C)C
Similarity: 0.945
Experimental value : 0.51
Predicted value : 1.376



Compound #5

CAS: N.A.
Dataset id:269 (Training Set)
SMILES: O=CC1CC(C)CC(C)(C)C1
Similarity: 0.941
Experimental value : 1.43
Predicted value : 1.06



Compound #6

CAS: N.A.
Dataset id:631 (Training Set)
SMILES: O=C1CCC(CC1)C(C)(C)CC
Similarity: 0.933
Experimental value : 1.45
Predicted value : 1.086

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 = 0.952 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.633 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✓	Concordance for similar molecules Concordance index = 0.557 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 = 0.939 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:

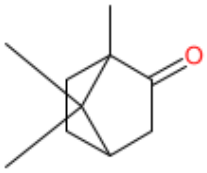
✓ The feature has a good assessment, model is reliable regarding this aspect.

⚠ 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 1.26 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
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a moderate 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)

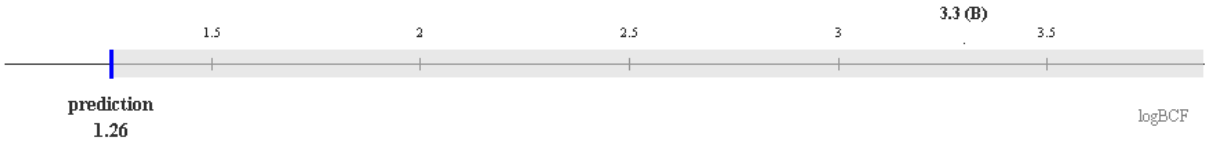
The following relevant fragments have been found: Carbonyl residue (SR 02); >C=O group (PG 09)

Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted BCF [log(L/kg)]: 1.26
Predicted BCF [L/kg]: 18
Predicted BCF from sub-model 1 (HM) [log(L/kg)]: 1.54
Predicted BCF from sub-model 2 (GA) [log(L/kg)]: 1.22
Predicted LogP (MLogP): 2.36
Structural Alerts: Carbonyl residue (SR 02); >C=O group (PG 09)
Reliability: The predicted compound is outside the Applicability Domain of the model
Remarks:
none

2. Possible Use and Uncertainty

Threshold 3.3 (bioaccumulative)

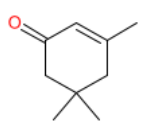
Following, a chart showing the predicted value together with its conservative confidence interval for safe classification.
For the threshold logBCF = 3.3, the current compound can not be associated (due to its Applicability Domain index value) to any conservative interval.
No safe classification can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

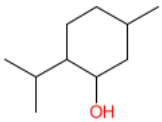
3.1 Applicability Domain:
Similar Compounds, with Predicted and Experimental Values



Compound #1

CAS: 78-59-1
Dataset id:477 (Test Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848
Experimental value : 0.14
Predicted value : 0.987

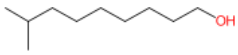
Alerts (found also in the target): Carbonyl residue (SR 02); >C=O group (PG 09)



Compound #2

CAS: 2216-51-5
Dataset id:119 (Training Set)
SMILES: OC1CC(C)CCC1(C(C)C)
Similarity: 0.827
Experimental value : 0.89
Predicted value : 1.563

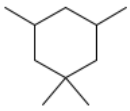
Alerts (not found also in the target): OH group (PG 06)



Compound #3

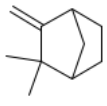
CAS: 25339-17-7
Dataset id:10 (Training Set)
SMILES: OCCCCCCCC(C)C
Similarity: 0.773
Experimental value : 2.51
Predicted value : 1.664

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



Compound #4

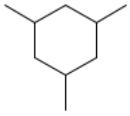
CAS: 50876-32-9
Dataset id:17 (Training Set)
SMILES: CC1CC(C)CC(C)(C)C1
Similarity: 0.771
Experimental value : 3.65
Predicted value : 3.135



Compound #5

CAS: 79-92-5
Dataset id:146 (Training Set)
SMILES: C=C2C1CCC(C1)C2(C)C
Similarity: 0.771
Experimental value : 2.98
Predicted value : 2.331

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



Compound #6

CAS: 1839-63-0
Dataset id:16 (Test Set)
SMILES: CC1CC(C)CC(C)C1
Similarity: 0.756
Experimental value : 3.34
Predicted value : 2.941

3.2 Applicability Domain:
Measured Applicability Domain Scores

✖	Global AD Index AD index = 0.712 Explanation: The predicted compound is outside the Applicability Domain of the model.
⚠	Similar molecules with known experimental value Similarity index = 0.837 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.76 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 0.745 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
⚠	Maximum error of prediction among similar molecules Max error index = 0.847 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 = 0.85 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.

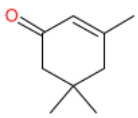
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

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

Fragment found: Carbonyl residue (SR 02)

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

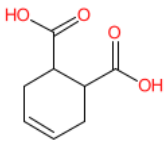
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 78-59-1
Dataset id:477 (Test Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848

Experimental value : 0.14
Predicted value : 0.987

Alerts (found also in the target): Carbonyl residue (SR 02); >C=O group (PG 09)

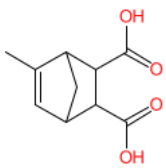


CAS: 88-98-2
Dataset id:121 (Training Set)
SMILES: O=C(O)C1CC=CCC1(C(=O)O)
Similarity: 0.654

Experimental value : 0.3
Predicted value : 0.107

Alerts (found also in the target): Carbonyl residue (SR 02)

Alerts (not found also in the target): COOH group (PG 01)



CAS: 50853-70-8
Dataset id:131 (Training Set)
SMILES: O=C(O)C2C1C=C(C)C(C1)C2(C(=O)O)
Similarity: 0.646

Experimental value : 0.74
Predicted value : 0.276

Alerts (found also in the target): Carbonyl residue (SR 02)

Alerts (not found also in the target): COOH group (PG 01)

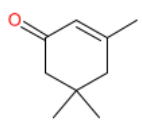
4.1 Reasoning:
Relevant Chemical Fragments and Moieties

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

Fragment found: >C=O group (PG 09)

This chemical contains a >C=O polar group. The presence of polar groups increases hydrophilicity, related to lower values of BCF.

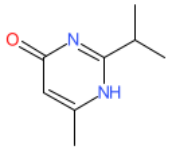
Following, the most similar compounds from the model's dataset having the same fragment.



CAS: 78-59-1
Dataset id:477 (Test Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848

Experimental value : 0.14
Predicted value : 0.987

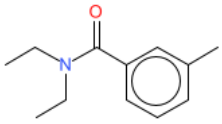
Alerts (found also in the target): Carbonyl residue (SR 02); >C=O group (PG 09)



CAS: 2814-20-2
Dataset id:455 (Test Set)
SMILES: O=C1N=C(NC(=C1)C)C(C)C
Similarity: 0.611

Experimental value : -0.38
Predicted value : 0.167

Alerts (found also in the target): Carbonyl residue (SR 02); >C=O group (PG 09)



CAS: 134-62-3
Dataset id:426 (Training Set)
SMILES: O=C(c1cccc(c1)C)N(CC)CC
Similarity: 0.598

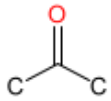
Experimental value : 0.38
Predicted value : 1.404

Alerts (found also in the target): Carbonyl residue (SR 02); >C=O group (PG 09)

Alerts (not found also in the target): Moiety (SMILES: O=Cc1ccccc1) (SR 01); 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:

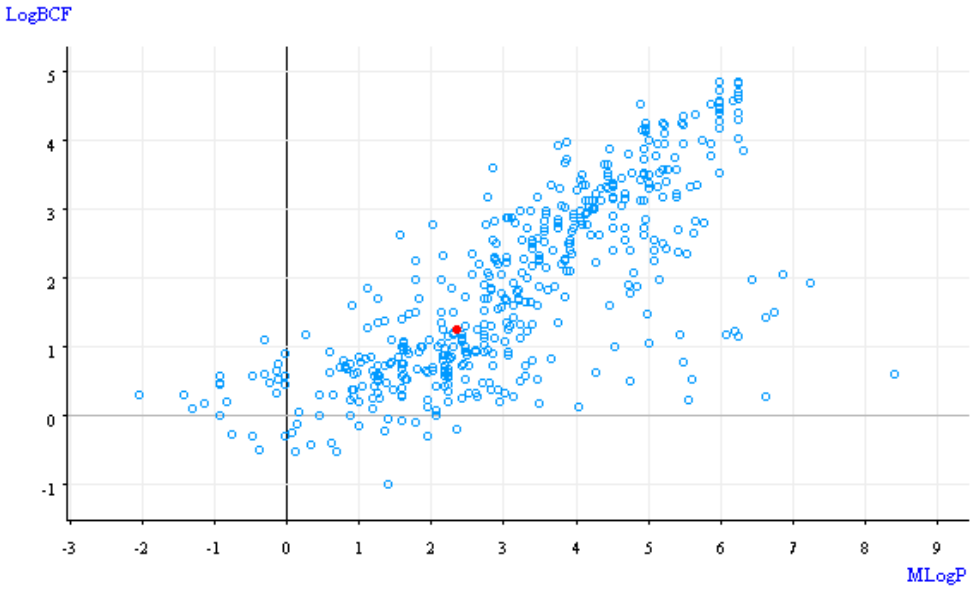


Fragment defined by the SMILES: CC(C)=O
The fragment has less than 3 occurrences in the model's training set

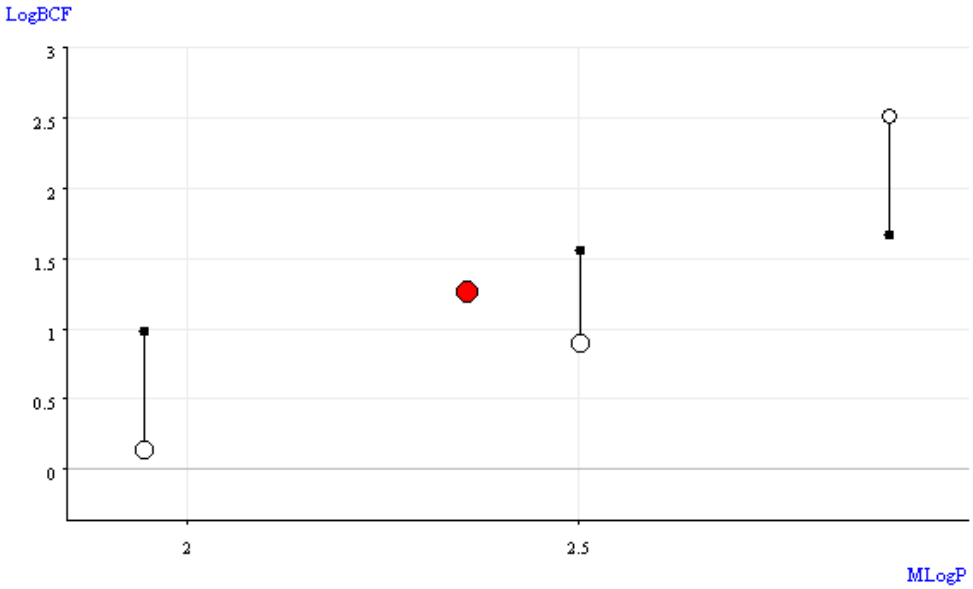
4.2 Reasoning:
Analysis of Molecular Descriptors

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.

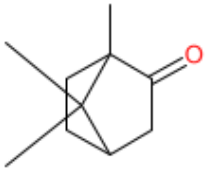


Following, a scatterplot of MLogP against response values only for 3 most similar compounds in the training set. Red dot is the value of the studied compound, black outlined circles represents experimental values of compounds from training set, black dots represents predicted value of the same compound;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



1. Prediction Summary

Prediction for compound Molecule 0 -

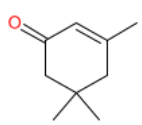


Prediction: Reliability:

Prediction is 1.36 log(L/kg), 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
- Accuracy of prediction for similar molecules found in the training set is not optimal
- some similar molecules found in the training set have experimental values that disagree with the predicted value
- the maximum error in prediction of similar molecules found in the training set has a moderate value, considering the experimental variability

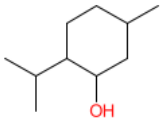
Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted BCF [log(L/kg)]: 1.36
Predicted BCF [L/kg]: 23
Predicted LogP (Meylan/Kowwin): 2.56
Predicted LogP reliability: Experimental
MW: 151.46
Ionic compound: no
Reliability: The predicted compound could be out of the Applicability Domain of the model
Remarks:
none

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



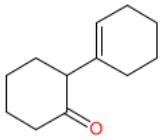
Compound #1

CAS: 78-59-1
Dataset id:119 (Training Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848
Experimental value : 0.3
Predicted value : 0.789



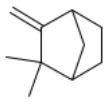
Compound #2

CAS: 1490-04-6
Dataset id:238 (Training Set)
SMILES: OC1CC(C)CCC1(C(C)C)
Similarity: 0.827
Experimental value : 1.18
Predicted value : 1.864



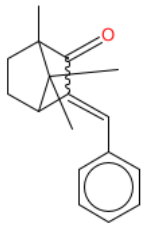
Compound #3

CAS: 1502-22-3
Dataset id:236 (Training Set)
SMILES: O=C1CCCCC1C2=CCCCC2
Similarity: 0.812
Experimental value : 0.7
Predicted value : 1.759



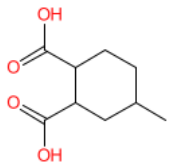
Compound #4

CAS: 79-92-5
Dataset id:347 (Training Set)
SMILES: C=C2C1CCC(C1)C2(C)C
Similarity: 0.771
Experimental value : 2.98
Predicted value : 2.451



Compound #5

CAS: 15087-24-8
Dataset id:422 (Training Set)
SMILES: O=C2C(=Cc1ccccc1)C3CCC2(C)C3(C)(C)
Similarity: 0.757
Experimental value : 2.5
Predicted value : 2.783



Compound #6

CAS: 57567-84-7
Dataset id:59 (Training Set)
SMILES: O=C(O)C1CCC(C)CC1(C(=O)O)
Similarity: 0.754
Experimental value : 0.38
Predicted value : 0.5

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 = 0.837 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.586 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
⚠	Concordance for similar molecules Concordance index = 0.616 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
⚠	Maximum error of prediction among similar molecules Max error index = 0.684 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 = 1 Explanation: reliability of logP value used by the model is good..
✓	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 = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

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

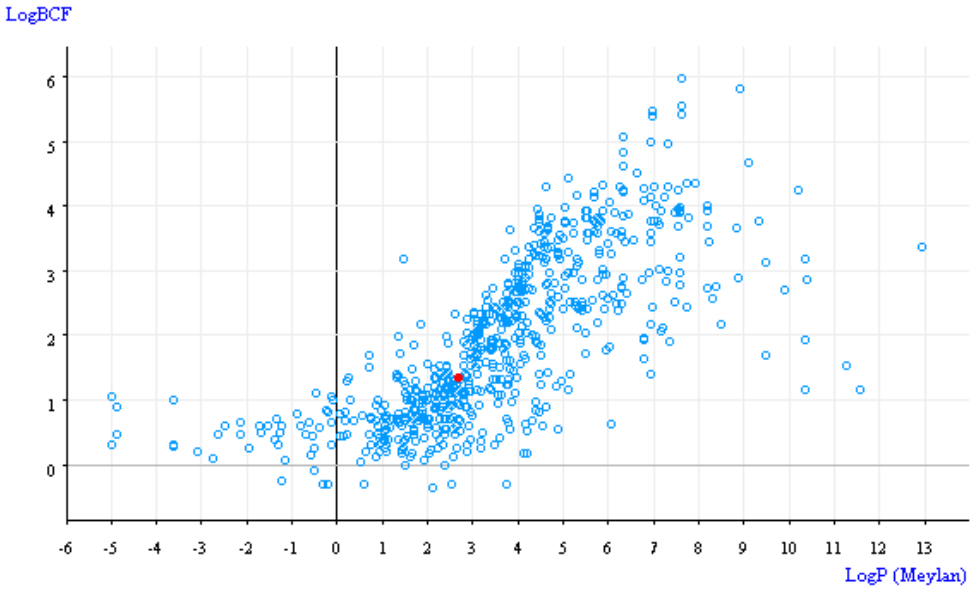
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

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

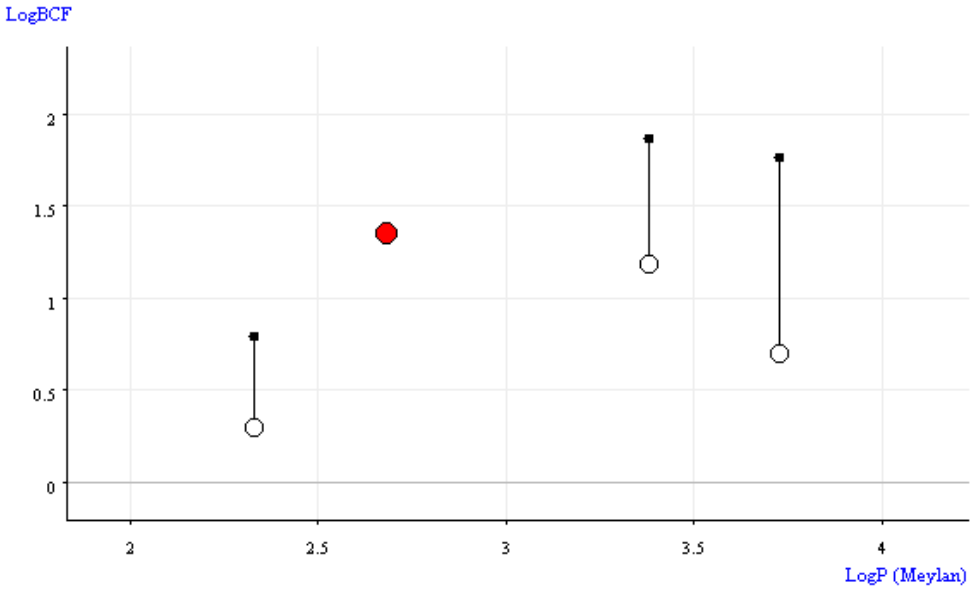
4.2 Reasoning:
Analysis of Molecular Descriptors

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

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

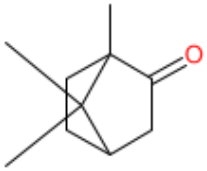


Following, a scatterplot of LogP (Meylan) against response values only for 3 most similar compounds in the training set. Red dot is the value of the studied compound, black outlined circles represents experimental values of compounds from training set, black dots represents predicted value of the same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



1. Prediction Summary

Prediction for compound Molecule 0 -

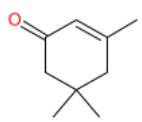


Prediction: Reliability:

Prediction is 1.52 log(L/kg), 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

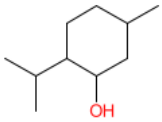
Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted BCF (up) [log(L/kg)]: 1.52
Predicted BCF (up) [L/kg]: 33
Predicted BCF (low) [log(L/kg)]: 1.33
Predicted BCF (low) [L/kg]: 21
Predicted BCF (mid) [log(L/kg)]: 1.37
Predicted BCF (mid) [L/kg]: 24
Predicted LogP (Meylan/Kowwin): 2.56
Predicted LogP reliability: Experimental
Predicted kM (Meylan): -0.23
Predicted kM reliability: Low
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



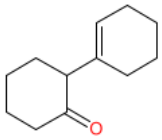
Compound #1

CAS: 78-59-1
Dataset id:462 (Training Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848
Experimental value : 0.67
Predicted value : 0.601



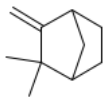
Compound #2

CAS: 2216-51-5
Dataset id:742 (Training Set)
SMILES: OC1CC(C)CCC1(C(C)C)
Similarity: 0.827
Experimental value : 0.89
Predicted value : 1.39



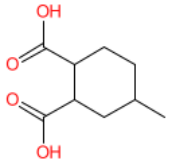
Compound #3

CAS: 1502-22-3
Dataset id:15 (Training Set)
SMILES: O=C1CCCCC1C2=CCCCC2
Similarity: 0.812
Experimental value : 0.49
Predicted value : 2.019



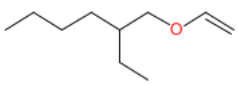
Compound #4

CAS: 79-92-5
Dataset id:136 (Training Set)
SMILES: C=C2C1CCC(C1)C2(C)C
Similarity: 0.771
Experimental value : 2.905
Predicted value : 3.068



Compound #5

CAS: 57567-84-7
Dataset id:115 (Training Set)
SMILES: O=C(O)C1CCC(C)CC1(C(=O)O)
Similarity: 0.754
Experimental value : -0.16
Predicted value : 0.507



Compound #6

CAS: 103-44-6
Dataset id:124 (Training Set)
SMILES: O(C=C)CC(CC)CCCC
Similarity: 0.751
Experimental value : 3.025
Predicted value : 2.738

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 = 0.837 Explanation: Only moderately similar compounds with known experimental value in the training set have been found..
✓	Accuracy of prediction for similar molecules Accuracy index = 0.284 Explanation: Accuracy of prediction for similar molecules found in the training set is good..
⚠	Concordance for similar molecules Concordance index = 0.742 Explanation: some similar molecules found in the training set have experimental values that disagree with the predicted value..
✓	Maximum error of prediction among similar molecules Max error index = 0.5 Explanation: the maximum error in prediction of similar molecules found in the training set has a low value, considering the experimental variability..
✓	Reliability of logP prediction LogP reliability = 1 Explanation: reliability of logP value used by the model is good..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

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

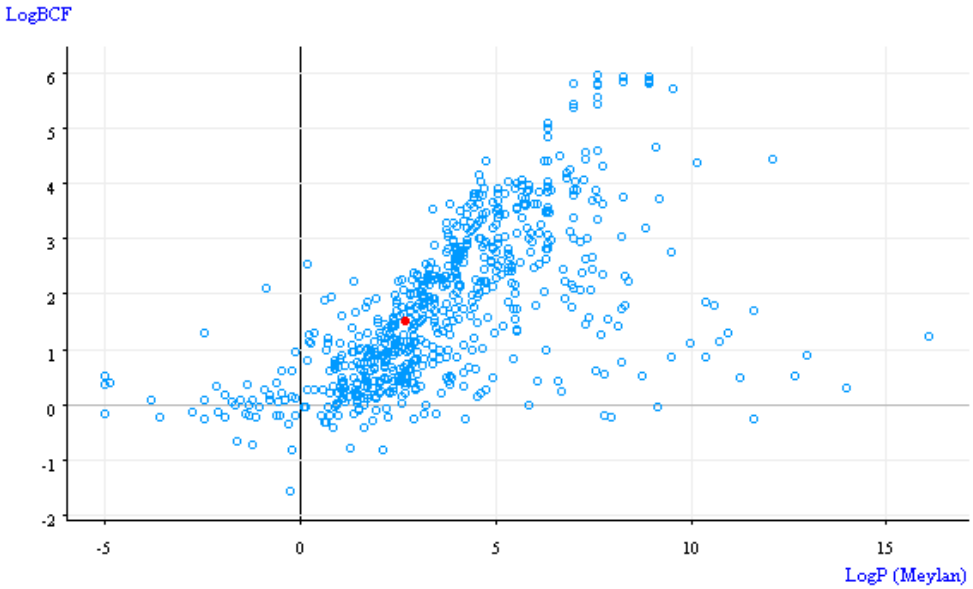
⚠ The feature has a non optimal assessment, this aspect should be reviewed by an expert.

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

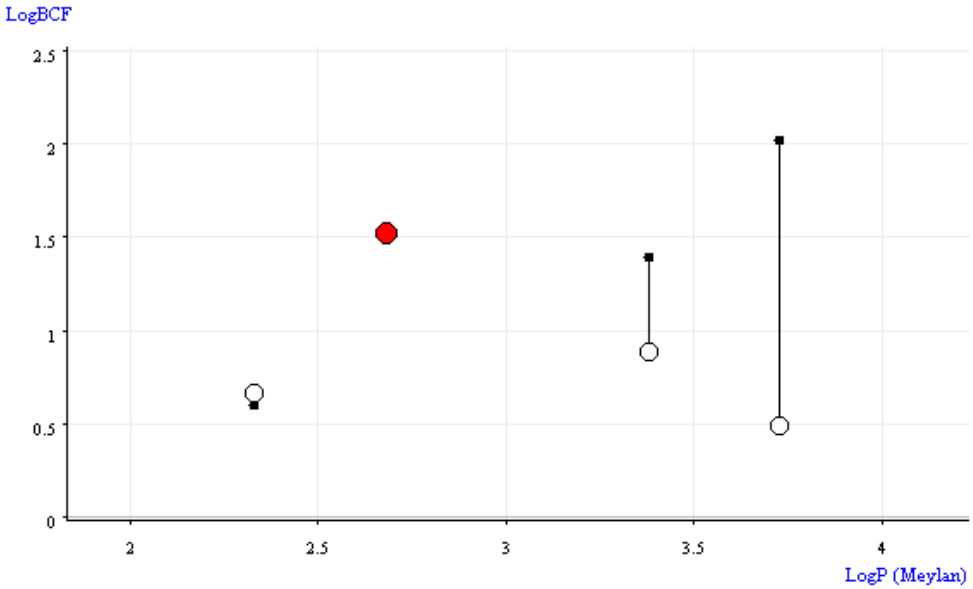
4.2 Reasoning:
Analysis of Molecular Descriptors

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

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

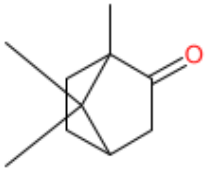


Following, a scatterplot of LogP (Meylan) against response values only for 3 most similar compounds in the training set. Red dot is the value of the studied compound, black outlined circles represents experimental values of compounds from training set, black dots represents predicted value of the same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



1. Prediction Summary

Prediction for compound Molecule 0 -



Prediction: Reliability:

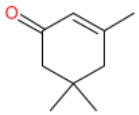
Prediction is 0.86 log(L/kg), but the result may be not reliable. A check of the information given in the following section should be done, paying particular attention to the following issues:
- Accuracy of prediction for similar molecules found in the training set is not optimal
- the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability

Compound: Molecule 0
Compound SMILES: O=C1CC2CCC1(C)C2(C)(C)
Experimental value: -
Predicted BCF [log(L/kg)]: 0.86
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

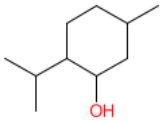
Compound #1

CAS: 2371-42-8
Dataset id:439 (Training Set)
SMILES: OC1(C)(CC2CCC1(C)C2(C)(C))
Similarity: 0.868
Experimental value : 1.477
Predicted value : 2.034



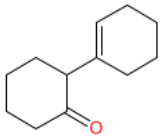
Compound #2

CAS: 78-59-1
Dataset id:51 (Training Set)
SMILES: O=C1C=C(C)CC(C)(C)C1
Similarity: 0.848
Experimental value : 0.39
Predicted value : 1.841



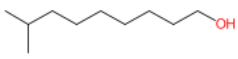
Compound #3

CAS: 2216-51-5
Dataset id:436 (Training Set)
SMILES: OC1CC(C)CCC1(C(C)C)
Similarity: 0.827
Experimental value : 0.892
Predicted value : 2.379



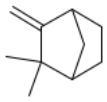
Compound #4

CAS: 1502-22-3
Dataset id:411 (Training Set)
SMILES: O=C1CCCCC1C2=CCCCC2
Similarity: 0.812
Experimental value : 0.59
Predicted value : 0.703



Compound #5

CAS: 25339-17-7
Dataset id:534 (Training Set)
SMILES: OCCCCCCCC(C)C
Similarity: 0.773
Experimental value : 2.51
Predicted value : 2.036



Compound #6

CAS: 79-92-5
Dataset id:58 (Training Set)
SMILES: C=C2C1CCC(C1)C2(C)C
Similarity: 0.771
Experimental value : 2.97
Predicted value : 3.198

3.2 Applicability Domain:
Measured Applicability Domain Scores

✗	Global AD Index AD index = 0.7 Explanation: The predicted compound is outside the Applicability Domain of the model.
✓	Similar molecules with known experimental value Similarity index = 0.836 Explanation: Strongly similar compounds with known experimental value in the training set have been ..
⚠	Accuracy of prediction for similar molecules Accuracy index = 0.902 Explanation: Accuracy of prediction for similar molecules found in the training set is not optimal..
✓	Concordance for similar molecules Concordance index = 0.347 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 = 1.487 Explanation: the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability..
✓	Atom Centered Fragments similarity check ACF index = 1 Explanation: all atom centered fragment of the compound have been found in the compounds of the training set..

Symbols explanation:

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

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

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

References and Documentation

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

Mutagenicity (Ames test) CONSENSUS model(version 1.0.4)

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

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

QSAR classification model for Mutagenicity (from CAESAR project)

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

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

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

QSAR classification model for Mutagenicity (SarPy/IRFMN)

References and Documentation

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

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

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

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

Carcinogenicity model (CAESAR)(version 2.1.10)

QSAR classification model for Carcinogenicity (from CAESAR project)

Carcinogenicity model (ISS)(version 1.0.3)

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

References and Documentation

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

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

Carcinogenicity model (IRFMN-Antares)(version 1.0.2)

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

Carcinogenicity oral classification model (IRFMN)(version 1.0.1)

Classification model for carcinogenicity (oral route).

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

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

References and Documentation

Carcinogenicity inhalation classification model (IRFMN)(version 1.0.1)

Classification model for carcinogenicity (inhalation route).

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

Quantitative model for the carcinogenicity inhalation route) Slope Factor.

Carcinogenicity in male rat (CORAL)(version 1.0.0)

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

Carcinogenicity in female Rat (CORAL)(version 1.0.0)

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

References and Documentation

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

KNN model for acute toxicity (LD50)

BCF model (CAESAR)(version 2.1.15)

QSAR regression model for fish BCF (from CAESAR project)

BCF model (Meylan)(version 1.0.4)

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

BCF model (Arnot-Gobas)(version 1.0.1)

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

References and Documentation

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

KNN (Read-Across) model for fish BCF