MaxTox Tutorial

Maxtox is a suite of models to predict toxicity of novel compounds based on their similarity to compounds with known toxicities. Compounds forming a model have the same Endpoint Toxicity.

For each model a set of compounds having the same toxicity endpoint are chosen. Substructures occurring (in more than 1 compound and which consist of more than 2 atoms) are extracted and dictionaries built. New compounds are compared to this dictionary and a fingerprint denoting the presence/absence (of dictionary fragments in the test molecule) is generated. This fingerprint is used in a statistical Random Forest Model (prebuilt and optimised) to generate a prediction of toxicity against the particular endpoint.

Maxtox is delivered as an open source application. It uses open source packages like CDK, Restlet, Jena & R to generate models and predict toxicities. It is part of the [OpenTox](#) initiative to create freely accessible resources to predict toxicity.

Maxtox can consume and respond in RDF datatypes and works with the [OpenTox](#) API specifiction. The Maxtox application can be used as a component of other prediciton use-cases hosted from other servers, as long as the data transactions are performed according to the [OpenTox](#) API.

BUILDING MaxTox PREDICTION MODEL AND PREDICTING TOXICITY

The workflow for building a toxicity prediction model using MaxTox is described in fig 1. MaxTox provides the online service to fingerprint their dataset using Maximum Common Substructure Search . Also the toxicity can be predicted online using MaxTox prediction models (See below for the details). Presently MaxTox does not offer services to generate prediction model online so the user has to chose from one of the three stored Random Forest models to predict the toxicity (dashed out box in fig 1). This tutorial covers the instructions on how to fingerprint/predict toxic compounds using MaxTox online service and through cURL calls.
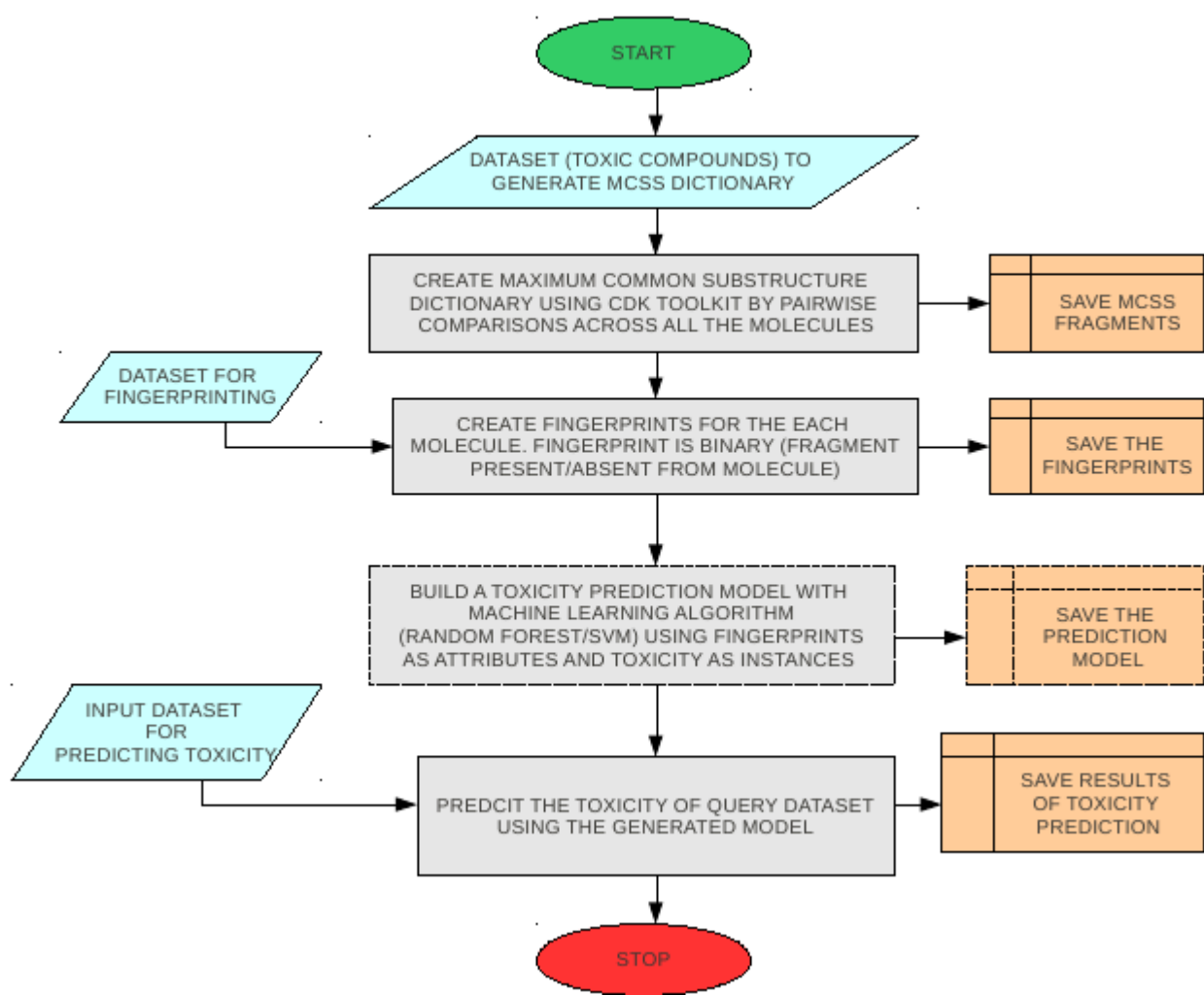
Fig 1. FlowChart for MaxTox Prediction Model building and Toxicity prediction

Create Fingerprints using MaxTox MCSS (Online service)

MaxTox MCSS utility can be used to generate MCSS (Maximum Common Substructure) dictionary for given set of compounds so that fingerprints for the query compound can be generated on the run using MaxTox algorithm and the generated MCSS dictionary. These fingerprints can then be fed to the models stored at MaxTox server to predict the toxicity of the query compound.

1. Submit the URI for the dataset whose MCSS dictionary has to be created using MaxTox at the MaxTox MCSS homepage at URL http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxMCSS/findMCSS

Fig 2. MaxTox MCSS home page

2. Once the dictionary generation task is submitted, it takes a while to complete the task. On completion of task, the result URI for the MCSS model generated is given.

Fig 3. Task details after completion of MCSS dictionary generation

3. Model details, dataset and fragment set URIs can be seen at detail page for models when the user clicks on the model URI for example, http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxMCSS/model/Model_c6ccd6d1_ef31_4f37_af18_7a57978f3bd4

Fig 4. MCSS Model details

To submit a set of compounds for fingerprint generation using the MCSS model built, click on the **"Fingerprint a dataset using this model"** tab.

4. If user wants to see the MCSS fragments generated from the given compounds, the SMILES for fragments are available on clicking the fragment set URI given on the model details page (Fig 4).

Fig 5. MCSS fragments generated using MCSS search (the given example contains only one MCSS fragment).

5. Submit the dataset URI whose fingerprints have to be generated and click **submit** button. The URL for fingerprinting : http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxMCSS/predict

Fig 6. Submitting a dataset URI for fingerprinting task.

If user wants to select some other model to create fingerprints, all other fingerprinting models are available at http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxMCSS/model .

6. Once the dataset is submitted to create fingerprints, MaxTox returns a task id stating the curent status and other details like date of start, date of end, description, percentage completed etc for the fingerprinting task initiated. It takes a while to complete the job, so refresh the page after a while.

Fig 7. Running a fingerprinting task

7. When the task is completed the status is set 'completed' and a URI is returned which gives a link to the result dataset uploaded on the AMBIT server.

Fig 8. Completed fingerprinting task

Create Fingerprints using MaxTox MCSS (using cURL)

Submit a dataset for creation of MCSS dictionary. The maximum Common Substructure fragments are extracted and a MCSS dictionary is created using CDK

```
curl -X POST -d "dataset_uri=http://apps.ideaconsult.net:8080/ambit2/dataset/272?max=2"
http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxMCSS/algorithm/MCSSFinder
```

Get the task details. Retrieve the model id from the details.

```
curl -X GET http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxMCSS/task/Task_ac9b584f_a28b_4931_8319_e31b8cb0e510
```

Get the details about MCSS dictionary model

```
curl -X GET http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxMCSS/model/Model_c6ccd6d1_ef31_4f37_af18_7a57978f3bd4
```

Submit the dataset for creating fingerprints

```

```

Get the result dataset URI

```

```

Predict Toxicity Using MaxTox (Online service)

1. Go to http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxTest/

Fig 9. Homepage for MaxTox

2. To predict toxicity of a compound, navigate to Predict Toxicity link on the left. This link leads to page for "**Submit a Prediction Task – step 1**" which currently contains three different Random Forests models generated using MCSS (Maximum Common Substructure Search) extracted from three different datasets.

Fig 10. Selecting a Model for toxicity prediction

Select one of the three models and click submit on the bottom of the page which follows to the page for " **Submit a Prediction Task – step 2**".

Detailed description of models about dataset used, endpoint and fragmentset is available at following URI, for example for model 1 :

http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxTest/model/1

Fig 11. Details of Model 1 submitted on MaxTox

3. Submit the URI of the dataset whose toxicity is to be predicted by MaxTox. Note that dataset or the compound has to be first loaded on AMBIT server (http://ambit.uni-plovdiv.bg:8080/ambit2),

for example, http://ambit.uni-plovdiv.bg:8080/ambit2/dataset/6?max=5

Completion of this step may take a while.

NOTE : As presently MaxTox has not implemented A&A, the dataset URI's may be provided from ideaconsult server, example of dataset URI :

http://apps.ideaconsult.net:8080/ambit2/dataset/272?max=2

Fig 12. Enter the URI of the dataset whose toxicity is to be predicted

Once the prediction task is submitted, MaxTox returns a task id which contains full task details like date, status, percentage completed, result URI etc.

Fig 13. Task Id for a running prediction task.

5. When the user comes back after a while and refreshes the task page or clicks on the task id, the result URI is displayed if the task is completed successfully or if the task is interrupted , an error message is displayed explaining the reason for task failure.

Fig 14. Completed Prediction task

The Result URI links to the prediction results for the submitted dataset stating whether the compounds are predicted to be toxic or not according to the MaxTox model used.

Predict Toxicity Using MaxTox (using cURL commands)

To get the list of models available on MaxTox

In text form :

curl -X GET -H "Accept:text" http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxTest/model

In rdf/xml form:

curl -X GET http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxTest/model

To submit a dataset for prediction using model 1

curl -X POST -H "Content-Type:application/x-www-form-urlencoded" -d "dataset_uri=http://apps.ideaconsult.net:8080/ambit2/dataset/272?max=2" http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxTest/model/1

Above command returns the prediction task Id, which can be used to retrieve the task status as follows

curl http://opentox2.informatik.uni-freiburg.de:8080/MaxtoxTest/task/Task_20110225121806

On completion of task, result dataset URI is returned which can be retrieved using the following curl command

curl -X GET http://apps.ideaconsult.net:8080/am63t2/dataset/19276