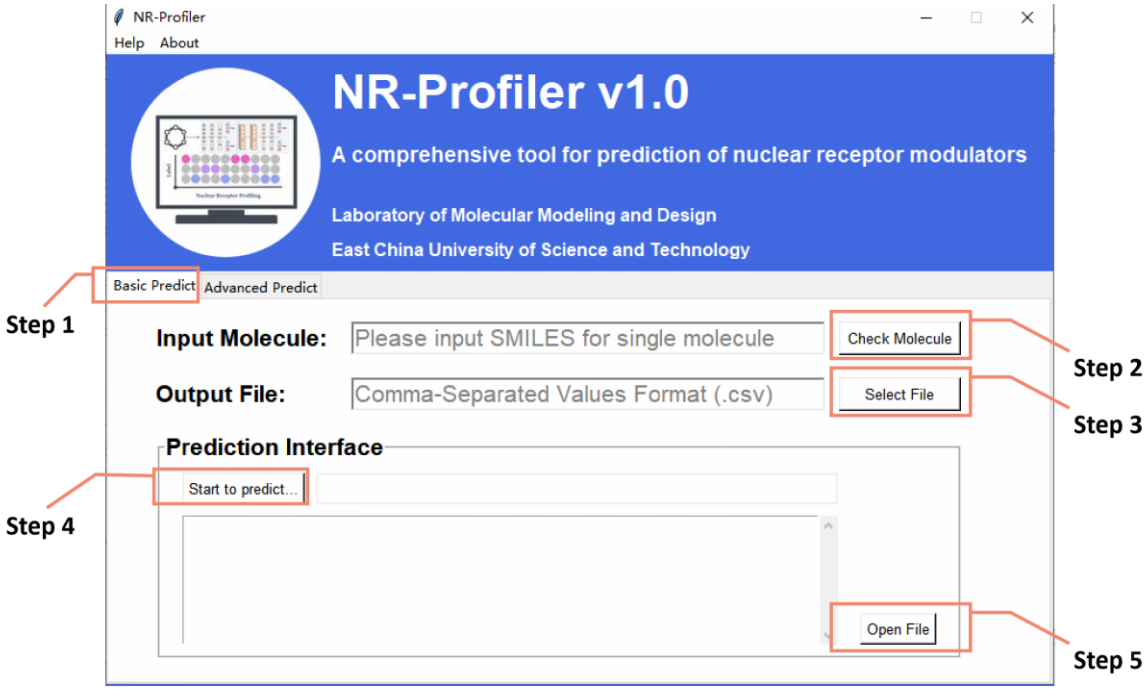
Documentation

The NR-Profiler (v1.0) is a software written in python package Tkinter (version 8.6.11), which can be used for prediction of nuclear receptor modulators. In this study, we used two deep learning algorithms to build the multi-task multi-classification models. In order to improve the predictive ability and robustness of a single estimator, we used the soft voting method to build the consensus model. With the software, researchers can input their in-house data and then obtain predictive lists. The help document is as below.

1. Basic Prediction



Step 1: The software depends on the JAVA environment. Please check if the Windows operating system has the JAVA environment. If the JAVA environment is not detected, you need to install the JAVA environment first. Only need to be checked once. The basic prediction type is to predict the nuclear receptor profiling for single-molecule.

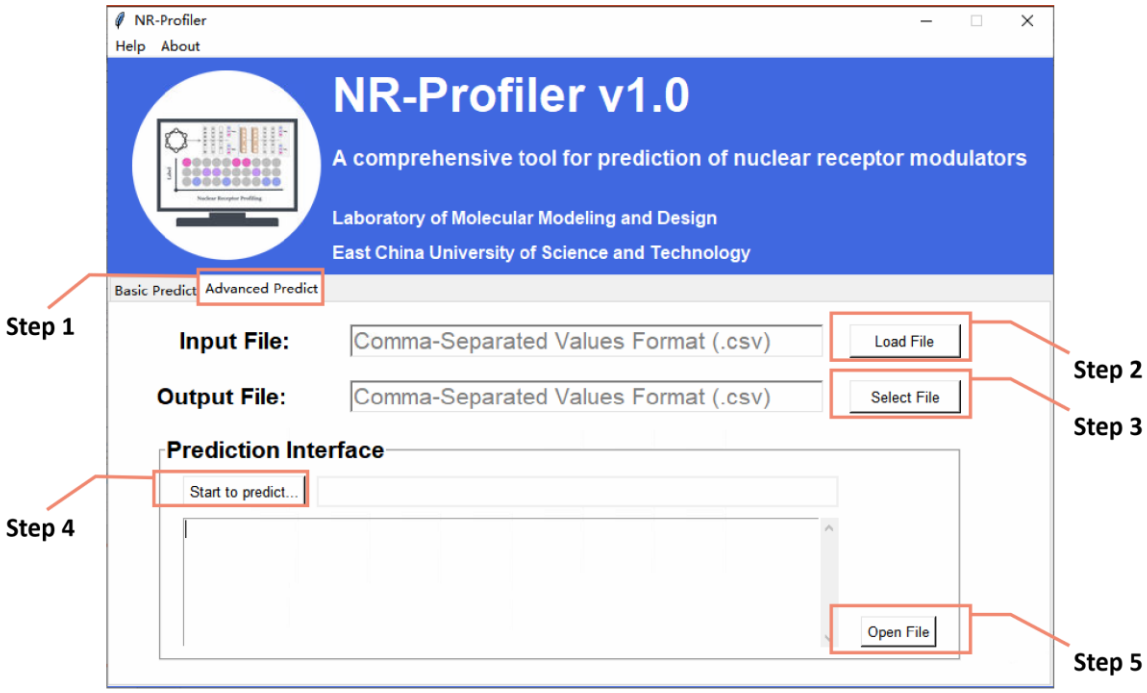
Step 2: Please input canonical SMILES for single-molecule and click "*Check Molecule*" button to ensure the correct molecular structure.

Step 3: Please select the output path of the prediction results and name the output file.

Step 4: Please click "*Start to predict*" button and wait a few seconds until the task view displays "Finish". It is estimated that the time required for the software to predict a compound is about 12 seconds on the Intel Core i5-11300H CPU (16 GB system memory).

Step 5 (Option): Users can click the "*Open File*" button to quickly open the output file.

2. Advanced Prediction



Step 1: The advanced prediction type is to predict the nuclear receptor profiling for thousands of compounds.

Step 2: Please click "*Load File*" button and select the input file. The input file must be a CSV format file and use "**SMILES**" as the header.

Step 3: Please select the output path of the prediction results and name the output file.

Step 4: Please click "*Start to predict*" button and wait for a while (100 molecules in about 1 minute) until the task view displays "Finish".

Step 5 (Option): Users can click the "*Open File*" button to quickly open the output file.

3. Guide

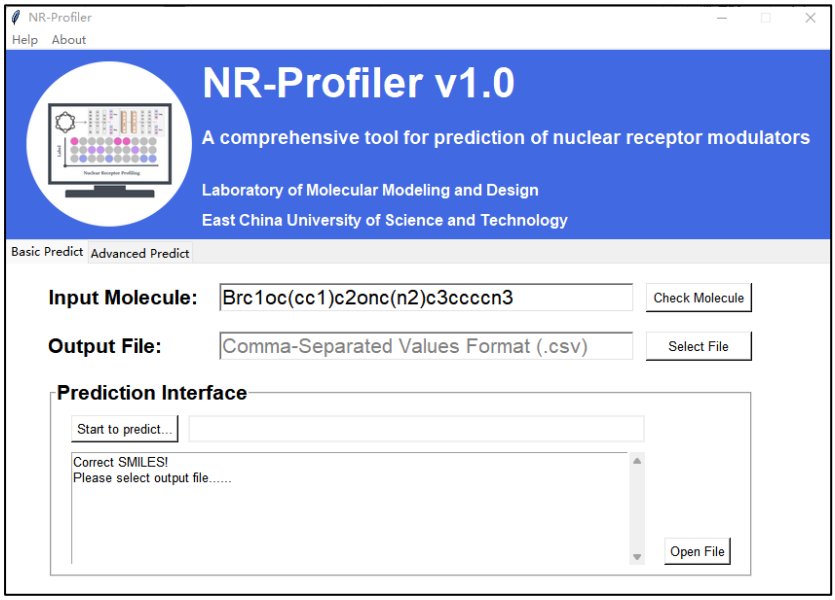
Users can choose two types, including basic and advanced predictions. Then, users can make predictions by simply inputting canonical SMILES format of the compound. When users select the advanced prediction type, the input file must be a CSV format file and use "**SMILES**" as the header. Finally, users can select the output path of the prediction results and start to predict. The predicted results are saved in a **CSV** format file. Users can choose these interactions with label = "Binder" for further experimental validation. It is estimated that the time required for the software to predict a compound is about 12 seconds on the Intel Core i5-11300H CPU (16 GB system memory). As the number of predicted compounds increases, the time it takes to predict will increase (100 molecules in about 1 minute). It depends on the hardware configuration of the computer.

4. Quick Reference

4.1 Basic Prediction

a) Input the SMILES of a compound, such as "Brc1oc(cc1)c2onc(n2)c3ccccn3"

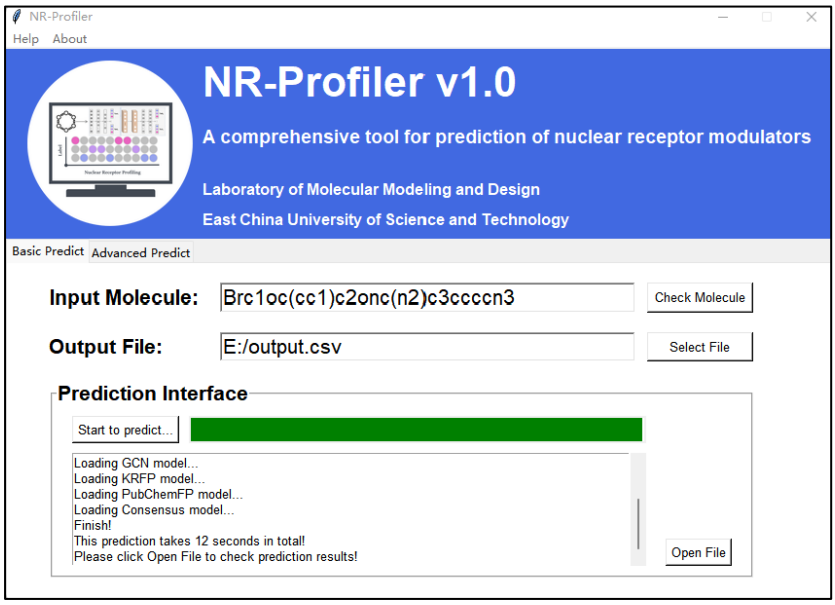
b) Click "*Check Molecule*" button to check if the SMILES is correct



c) Click "*Select File*" button to select the output file (CSV format)

d) Click "*Start to predict*" button to start loading the model and wait for the prediction to end

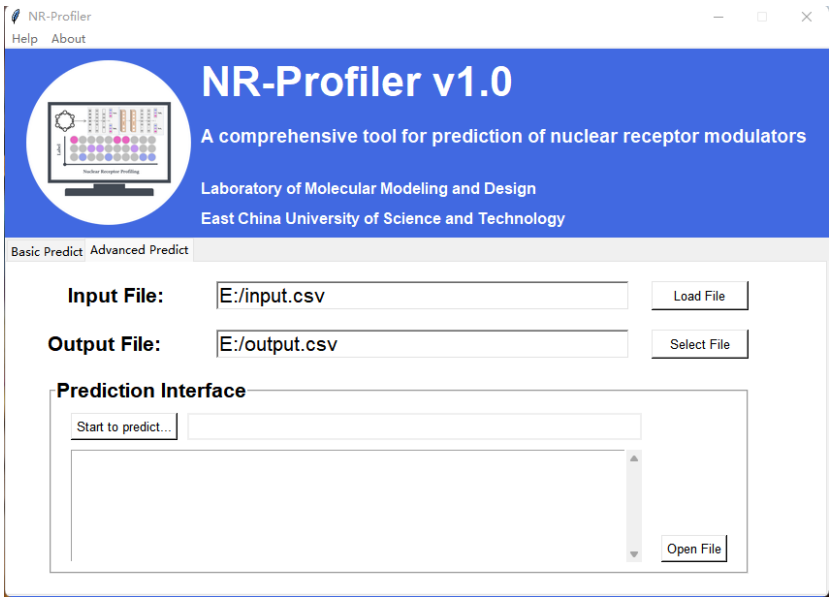
e) Click "*Open File*" button to open the output file and analyze the prediction results

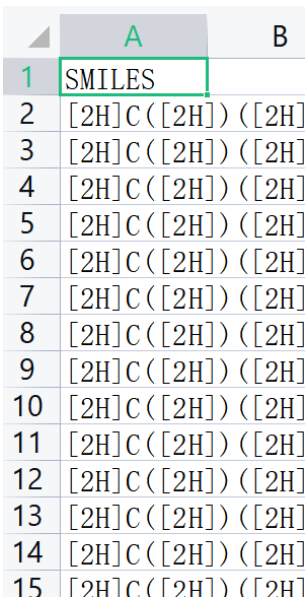


4.2 Advanced Prediction

a) Click "*Load File*" button and select the input file. The input file must be a **CSV** format file and use "**SMILES**" as the header.

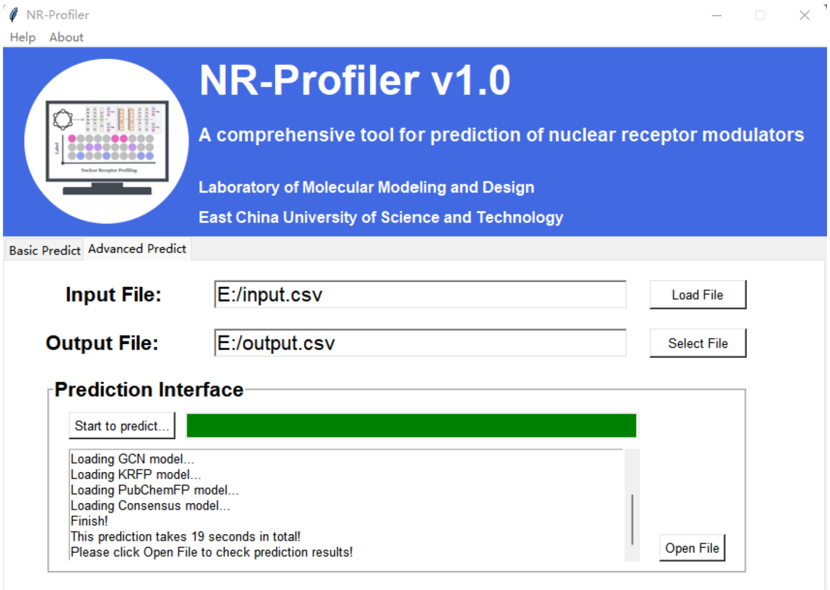
b) Click "*Select File*" button to select the output file (**CSV** format)



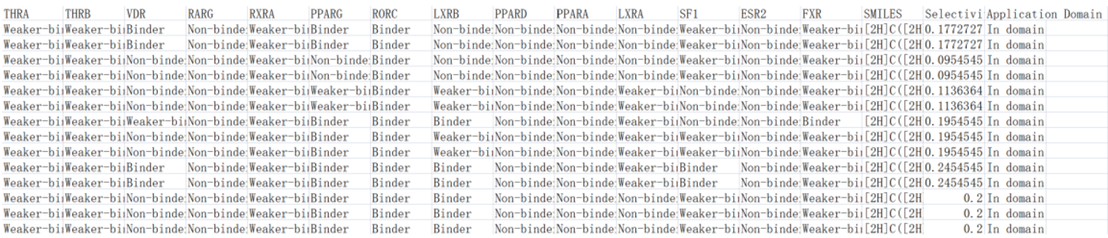


c) Click "*Start to predict*" button to start loading the model and wait for the prediction to end

d) Click "*Open File*" button to open the output file and analyze the prediction results



e) Prediction Result Display



THRA	THRB	VDR	RARG	RXRA	PPARG	RORC	LXRB	PPARD	PPARA	LXRA	SF1	ESR2	FXR	SMILES	Selectivi	Application Domain
Weaker-bi	Weaker-bi	Binder	Non-binde	Weaker-bi	Binder	Binder	Non-binde	Non-binde	Non-binde	Non-binde	Weaker-bi	Non-binde	Weaker-bi	[2H]C([2H	0.1772727	In domain
Weaker-bi	Weaker-bi	Binder	Non-binde	Weaker-bi	Binder	Binder	Non-binde	Non-binde	Non-binde	Non-binde	Weaker-bi	Non-binde	Weaker-bi	[2H]C([2H	0.1772727	In domain
Weaker-bi	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Non-binde	Binder	Non-binde	Non-binde	Non-binde	Non-binde	Weaker-bi	Non-binde	Weaker-bi	[2H]C([2H	0.0954545	In domain
Weaker-bi	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Non-binde	Binder	Non-binde	Non-binde	Non-binde	Non-binde	Weaker-bi	Non-binde	Weaker-bi	[2H]C([2H	0.0954545	In domain
Weaker-bi	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Weaker-bi	Binder	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Non-binde	Non-binde	Weaker-bi	[2H]C([2H	0.1136364	In domain
Weaker-bi	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Weaker-bi	Binder	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Non-binde	Non-binde	Weaker-bi	[2H]C([2H	0.1136364	In domain
Weaker-bi	Weaker-bi	Weaker-bi	Non-binde	Weaker-bi	Binder	Binder	Binder	Non-binde	Non-binde	Weaker-bi	Non-binde	Non-binde	Binder	[2H]C([2H	0.1954545	In domain
Weaker-bi	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Binder	Binder	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Weaker-bi	Non-binde	Weaker-bi	[2H]C([2H	0.1954545	In domain
Weaker-bi	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Binder	Binder	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Weaker-bi	Non-binde	Weaker-bi	[2H]C([2H	0.1954545	In domain
Weaker-bi	Weaker-bi	Binder	Non-binde	Weaker-bi	Binder	Binder	Binder	Non-binde	Non-binde	Weaker-bi	Binder	Non-binde	Weaker-bi	[2H]C([2H	0.2454545	In domain
Weaker-bi	Weaker-bi	Binder	Non-binde	Weaker-bi	Binder	Binder	Binder	Non-binde	Non-binde	Weaker-bi	Binder	Non-binde	Weaker-bi	[2H]C([2H	0.2454545	In domain
Weaker-bi	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Binder	Binder	Binder	Non-binde	Non-binde	Non-binde	Weaker-bi	Non-binde	Weaker-bi	[2H]C([2H	0.2	In domain
Weaker-bi	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Binder	Binder	Binder	Non-binde	Non-binde	Non-binde	Weaker-bi	Non-binde	Weaker-bi	[2H]C([2H	0.2	In domain
Weaker-bi	Weaker-bi	Non-binde	Non-binde	Weaker-bi	Binder	Binder	Binder	Non-binde	Non-binde	Non-binde	Weaker-bi	Non-binde	Weaker-bi	[2H]C([2H	0.2	In domain

In the output file, **the title of the table** is the name of each nuclear receptor, the SMILES of each compound, Selectivity Score and Application Domain. Users can choose these interactions with label = "**Binder**" and lower **Selectivity Score**, compounds within **Application Domain** for further experimental validation.

5. Developer Information

Maintainer: Dr. Lou, Dr. Wang, Dr. Wu and Professor Tang

Contact US: ytang234@ecust.edu.cn

Laboratory of Molecular Modeling and Design

School of Pharmacy, East China University of Science and Technology