Instructions for running a model on a local machine

JADBIO allows the user to download a model and run it on a local machine.

To run our model locally, the user must meet the following requirements:

1. Java SE Development Kit version 17 (https://www.oracle.com/java/technologies/javase/jdk17-archive-downloads.html)

2. the Java executor (contained in model.zip - filename: jadbio-1.4.55-model-exe.jar)

3. created models based on enzymatic reactions, metabolic pathways and predicted microbial metabolites (contained in model.zip - filenames: ecs.bin, pathways.bin, PredictedMetabolites.bin)

After installing Java SE JDK, model.zip must be saved somewhere on the local machine. After saving model.zip, provided by the authors, the folder must be extracted (e.g. with WinZip, 7zip). The model must be executed with the command prompt (cmd) (Fig. 1, 2).

Step 1

Using the `cd path` command (Fig. 1), the user navigates to the same directory (e.g. Folder) that contains the model executor (.jar) and the model (.bin).

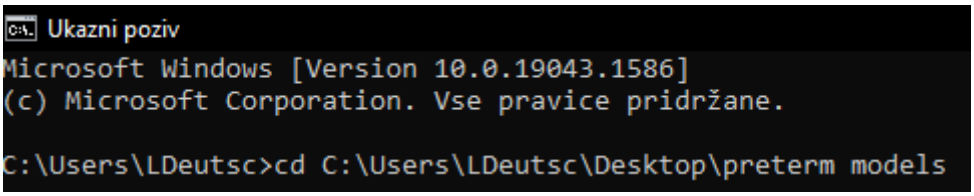

Figure 1. First command to navigate to the folder containing the model. In this case, we used the `cd C:\Users\LDeutsc\Desktop\preterm models` command because Executor and Models were in the preterm models folder on the desktop.

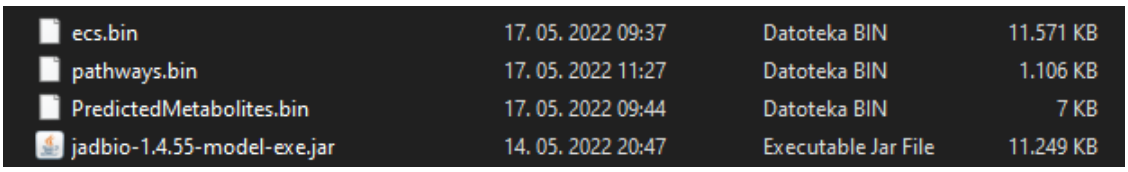

Figure 2. The two files needed for the overview of the model. Both files are contained in the model.zip folder.

Step 2

The next step is to preview the model using the following command:

`Java --enable-preview -jar jadbio-1.4.55-model-exe.jar -n <NAME OF THE MODEL>.bin`

This allows the user to get an overview of the model, key features and information about the analysis (which algorithm was used, version of JADBIO and other information about the model) (Fig. 3).

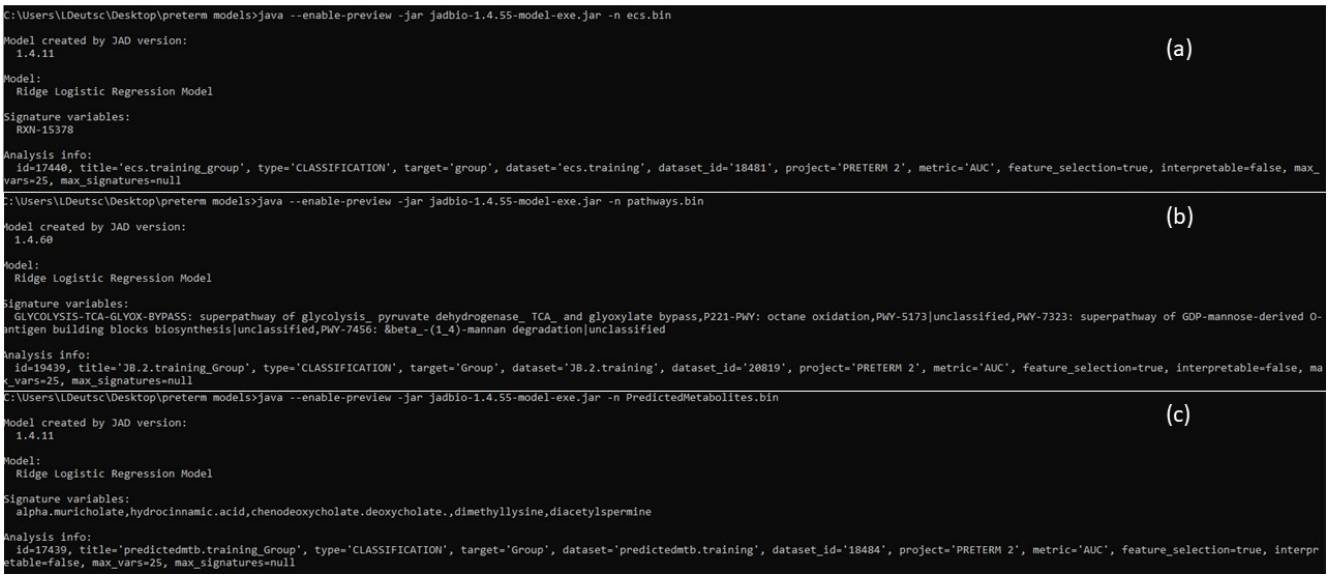

Figure 3. Overview of the models for (a) enzymatic reactions, (b) metabolic pathways and (c) predicted metabolites.

Step 3

To test the model, the user must prepare data. For the model to work properly, the output of HUMAnN 3.0 and MelonnPan must be used. The user needs to prepare the dataset as shown in Figure 4 and save it in a .csv document (comma separated values). Names of features must be prepared in a way, that is presented in preview of the model (section named Signature variables). After preparation, the prepared dataset must be saved in the same directory as the model executor and the model itself.

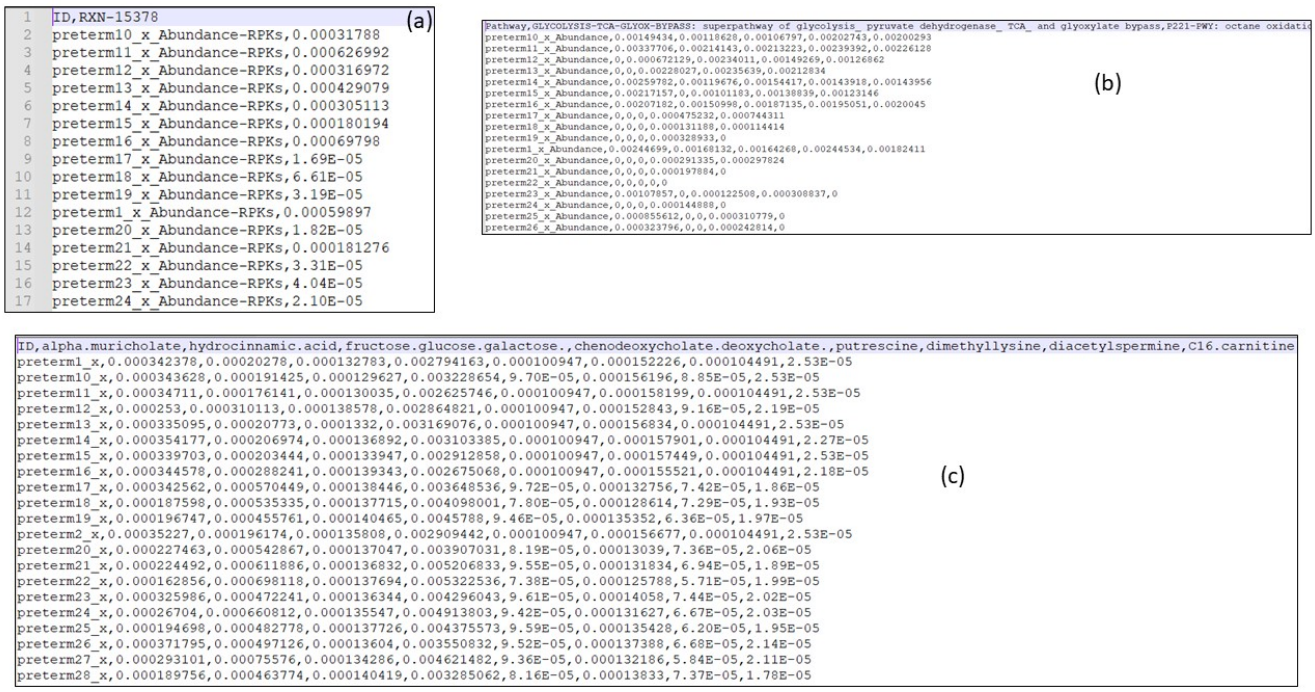

Figure 4. Prepared datasets for enzymatic reactions (a), metabolic pathways (b) and predicted metabolites (c).

Step 4

After saving the dataset, the user must use the next command in the terminal:

```
java --enable-preview -jar jadbio-1.4.55-model-exe.jar -m <NAME OF THE MODEL>.bin -i <NAME OF PREPARED DATASET>.csv -o result.csv
```

This command runs the model on test data (<NAME OF PREPARED DATASET>.csv in our case) and creates a new dataset with predictions (result.csv) (Fig. 5).

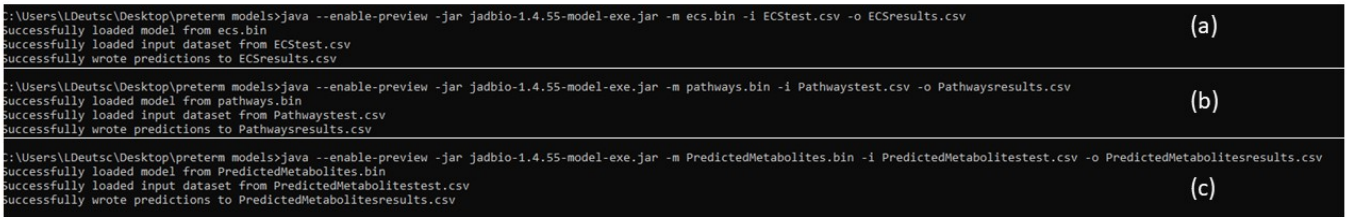

Figure 5. Executing the model and creating the result .csv file with predictions in the same directory based on a given enzymatic reactions (a), metabolic pathways (b) and predicted metabolites (c).

Step 5

After model execution, the user can check the calculated predictions by opening the .csv file directly by clicking on the created .csv file and opening it in any data analysis program (Excel, Past, R ...). As shown in Fig. 6 the model classifies the data between two groups (preterm and control). The first column is the same as in the test data created by the user.

(a)

	A	B	C
1	Sample na	Prob (class = Control)	Prob (class = Preterm)
2	preterm10	0.332290189	0.667709811
3	preterm11	0.022846072	0.977153928
4	preterm12	0.33428622	0.66571378
5	preterm13	0.14210351	0.85789649
6	preterm14	0.360881065	0.639118935
7	preterm15	0.660219252	0.339780748
8	preterm16	0.0114512	0.9885488
9	preterm17	0.907183957	0.092816043
10	preterm18	0.857293973	0.142706027
11	preterm19	0.893911389	0.106088611

(b)

	A	B	C
1	Sample na	Prob (class = Control)	Prob (class = Preterm)
2	preterm10	0.439365634	0.560634366
3	preterm11	0.311599427	0.688400573
4	preterm12	0.510187055	0.489812945
5	preterm13	0.481385161	0.518614839
6	preterm14	0.437772542	0.562227458
7	preterm15	0.521772789	0.478227211
8	preterm16	0.394080668	0.605919332
9	preterm17	0.64627694	0.35372306

(c)

	A	B	C
1	Sample na	Prob (class = Control)	Prob (class = Preterm)
2	preterm1_	0.002657309	0.997342691
3	preterm10	0.002889081	0.997110919
4	preterm11	2.28E-05	0.999977227
5	preterm12	0.958998991	0.041001009
6	preterm13	9.09E-04	0.999090718
7	preterm14	1.49E-04	0.999851037
8	preterm15	2.34E-04	0.999765739
9	preterm16	0.003400977	0.996599023
10	preterm17	0.999999999	1.01E-09
11	preterm18	1	1.85E-14
12	preterm19	1	1.53E-12

Figure 6. The newly created .csv file with predictions calculated from enzymatic reactions (a), metabolic pathways (b) and predicted metabolites (c).