CLAIMS

1. A method for constructing a data processing model, comprising:

acquiring model description parameters and sample data of a target data processing model, wherein the model description parameters are provided by a user for the target data processing model and comprise a name of the target data processing model, a data label obtainable by the target data processing model, a data feature useable by the target data processing model, and an evaluation metric of the target data processing model;

determining a base model according to the model description parameters and the sample data; and

training the base model according to the sample data to obtain the target data processing model,

wherein determining the base model according to the model description parameters and the sample data comprises:

determining a model type of the base model from a preset model type group according to the data label, wherein the preset model type group comprises at least a classification type and a regression prediction type;

selecting a plurality of models from models meeting the model type of the base model to serve as a plurality of candidate models;

establishing a pre-trained model for each of the plurality of candidate models according to the sample data and the data feature, and determining a model parameter of each pre-trained model;

determining an adaptability of each pre-trained model according to the evaluation metric, wherein the adaptability of a pre-trained model is configured to represent a degree of applicability of data processing on the sample data based on a pre-trained model and a model parameter corresponding to the pre-trained model; and

determining the base model from the plurality of candidate models according to the adaptabilities of the pre-trained models.

2. The method of claim 1, wherein selecting a plurality of models from models meeting the model type of the base model to serve as candidate models comprises:

processing the sample data by using data processing models of which applied base models meet the model type of the base model in a preset data processing model library to obtain data processing results of the data processing models;

conducting statistics on evaluation metrics of the data processing models according to the data processing results of the data processing models;

determining a data processing model in which a degree of difference between its evaluation metric and the evaluation metric of the target data processing model is less than a preset degree threshold as a recommended model; and

determining a base model to which the recommended model is applied as a candidate model.

3. The method of claim 1, wherein after determining a model type of the base model from a preset model type group according to the data label, the method further comprises:

determining a data type of the sample data, wherein the data type of the sample data comprises at least one of the following: picture data and text data; and

selecting, according to the data type of the sample data and through filtering the model types of base models, a model type matching the data type of the sample data to serve as the model type of the base model.

4. The method of claim 1, wherein establishing a pre-trained model for each of the plurality of candidate models according to the sample data and the data feature comprises:

extracting a preset proportion of sample data from the sample data to serve as test data;

pre-processing the test data to obtain processed test data, wherein the pre-processing comprises: dimension reduction processing and/or cleaning processing;

selecting feature data from the processed test data according to the data feature; and

training each of the plurality of candidate models respectively according to the feature data to obtain the pre-trained model.

5. The method of claim 1, wherein determining an adaptability of each pre-trained model according to the evaluation metric comprises:

performing a parameter search and a structure search according to the pre-trained model and a corresponding model parameter to obtain a search result; and

determining the adaptability of the pre-trained model according to the search result and the evaluation metric.

6. The method of claim 5, wherein performing a parameter search and a structure search according to the pre-trained model and a corresponding model parameter to obtain a search result comprises:

performing the parameter search and the structure search through at least one of the following algorithms: a Monte Carlo algorithm, a grid algorithm, a Bayesian algorithm, and a genetic algorithm.

7. The method of claim 1, wherein determining the base model from the plurality of candidate models according to the adaptabilities of the pre-trained models comprises:

determining a model used by the pre-trained model with the highest adaptability as the base model.

8. The method of claim 7, wherein determining the base model from the plurality of candidate models according to the adaptabilities of the pre-trained models further comprises:

combining models used by the plurality of pre-trained models whose adaptabilities are greater than a preset adaptability threshold to obtain a combined model; and determining the combined model as the base model.

9. The method of claim 1, wherein the model description parameter further comprises: a model type of the base model, and a split rule of the sample data.

10. A data processing method, comprising:

acquiring to-be-processed target data; and

invoking a target data processing model to process the target data, wherein the target data processing model is obtained by determining a base model according to model description parameters and sample data of the target data processing model that are acquired and then training the base model according to the sample data,

wherein the model description parameters are provided by a user for the target data processing model and comprise a name of the target data processing model, a data label obtainable by the target data processing model, a data feature useable by the target data processing model, and an evaluation metric of the target data processing model, and

wherein the base model is determined based on the following:

determining a model type of the base model from a preset model type group according to the data label, wherein the preset model type group comprises at least a classification type and a regression prediction type;

selecting a plurality of models from models meeting the model type of the base model to serve as a plurality of candidate models;

establishing a pre-trained model for each of the plurality of candidate models according to the sample data and the data feature, and determining a model parameter of each pre-trained model;

determining an adaptability of each pre-trained model according to the evaluation metric, wherein the adaptability of a pre-trained model is configured to represent a degree of applicability of data processing on the sample data based on a pre-trained model and a model parameter corresponding to the pre-trained model; and

determining the base model from the plurality of candidate models according to the adaptabilities of the pre-trained models.

11. A data processing method, comprising:

displaying a data input interface; and

receiving model description parameters and sample data of a target data processing model input by a user on the data input interface, wherein the model description parameters and the sample data are configured to determine a base model, and the base model is configured to establish the target data processing model,

wherein the model description parameters comprise a name of the target data processing model, a data label obtainable by the target data processing model, a data feature useable by the target data processing model, and an evaluation metric of the target data processing model, and

wherein the base model is determined based on the following:

determining a model type of the base model from a preset model type group according to the data label, wherein the preset model type group comprises at least a classification type and a regression prediction type;

selecting a plurality of models from models meeting the model type of the base model to serve as a plurality of candidate models;

establishing a pre-trained model for each of the plurality of candidate models according to the sample data and the data feature, and determining a model parameter of each pre-trained model;

determining an adaptability of each pre-trained model according to the evaluation metric, wherein the adaptability of a pre-trained model is configured to represent a degree of applicability of data processing on the sample data based on a pre-trained model and a model parameter corresponding to the pre-trained model; and

determining the base model from the plurality of candidate models according to the adaptabilities of the pre-trained models.

12. The method of claim 11, wherein after receiving a model description parameter and sample data of a target data processing model input by a user on the data input interface, the method further comprises:

displaying, in response to a display operation, an execution file generated based on the model description parameter; and

receiving a modification operation on the execution file, and updating the execution file according to the modification operation.

13. An apparatus for constructing a data processing model, comprising:

an acquisition module configured to acquire model description parameters and sample data of a target data processing model, wherein the model description parameters are provided by a user for the target data processing model and comprise a name of the target data processing model, a data label obtainable by the target data processing model, a data feature useable by the target data processing model, and an evaluation metric of the target data processing model;

a determination module configured to determine a base model according to the model description parameter and the sample data; and

a training module configured to train the base model according to the sample data to obtain the target data processing model,

wherein the determination module comprises:

a first determination unit configured to determine a model type of the base model from a preset model type group according to the data label, wherein the preset model type group comprises at least a classification type and a regression prediction type;

a selection unit configured to select a plurality of models from models meeting the model type of the base model to serve as candidate models;

an establishment unit configured to establish a pre-trained model for each of the plurality of candidate models according to the sample data and the data feature, and determine a model parameter of each pre-trained model;

a second determination unit configured to determine an adaptability of each pre-trained model according to the evaluation metric, wherein the adaptability is configured to represent a degree of

applicability of data processing on the sample data based on a pre-trained model and a model parameter corresponding to the pre-trained model; and

a third determination unit configured to determine the base model from the plurality of candidate models according to the adaptabilities of the pre-trained models.

14. The apparatus of claim 13, wherein the selection unit comprises:

a processing sub-unit configured to process the sample data by using data processing models of which applied base models meet the model type of the base model in a preset data processing model library to obtain data processing results of the data processing models;

a statistics sub-unit configured to conduct statistics on evaluation metrics of the data processing models according to the data processing results of the data processing models;

a first determination sub-unit configured to determine a data processing model in which a degree of difference between its evaluation metric and the evaluation metric of the target data processing model is less than a preset degree threshold as a recommended model; and

a second determination sub-unit configured to determine a base model to which the recommended model is applied as a candidate model.

15. The apparatus of claim 13, wherein the establishment unit comprises:

an extraction sub-unit configured to extract a preset proportion of sample data from the sample data to serve as test data;

a pre-processing sub-unit configured to pre-process the test data to obtain processed test data, wherein the pre-processing comprises: dimension reduction processing and/or cleaning processing;

a selection sub-unit configured to select feature data from the processed test data according to the data feature; and

a training sub-unit configured to train each of the plurality of candidate models respectively according to the feature data to obtain the pre-trained model.

16. The apparatus of claim 13, wherein the third determination unit comprises: a third determination sub-unit configured to determine a model used by the pre-trained model with the highest adaptability as the base model.

17. The apparatus of claim 16, wherein the third determination unit further comprises: a fourth determination sub-unit configured to combine models used by the plurality of pre-trained models whose adaptabilities are greater than a preset adaptability threshold to obtain a combined model; and determine the combined model as the base model.

18. The apparatus of claim 13, wherein the model description parameter further comprises: a model type of the base model and a split rule of the sample data.

19. A server, comprising a processor and a memory for storing processor-executable instructions, wherein the steps of the method of any one of claims 1 to 9 are implemented when the processor executes the instructions.

20. A user terminal, comprising a processor and a memory for storing processor-executable instructions, wherein the steps of the method of any one of claims 10 to 12 are implemented when the processor executes the instructions.

21. A computer-readable storage medium, storing computer instructions thereon, wherein the steps of the method of any one of claims 1 to 9 are implemented when the instructions are executed.