

CLAIMS

1. A data type recognition method for recognizing data as first-type data or second-type data, wherein the method comprises:

acquiring data to be recognized, and using a preset anomaly detection model to detect whether the data to be recognized is first-type data; and

inputting other data than the first-type data recognized by the anomaly detection model, into a classification model for recognition, wherein the classification model classifies the other data as first-type data and second-type data,

wherein the anomaly detection model is obtained by pre-training based on a first sample data set, and the classification model is obtained by pre-training based on a second sample data set different from the first sample; and

the pre-training of the classification model comprises:

detecting, by the anomaly detection model, an abnormal sample data set from the second sample data set; and

using the abnormal sample data set to train the classification model.

2. The method according to claim 1, wherein the anomaly detection model is obtained by pre-training in the following method:

acquiring the first sample data set, wherein an amount of first-type data in the first sample data set is greater than that of second-type data; and

using the first sample data set to train the anomaly detection model.

3. The method according to claim 1, before using the abnormal sample data set to train the classification model, the method further comprising:

optimizing the abnormal sample data set based on a feature optimization algorithm.

4. A model training method for training an anomaly detection model and a classification model,

the anomaly detection model being used for detecting first-type data, and

the classification model being used for classifying other data than the first-type data detected by the anomaly detection model,

wherein the method comprises:

training the anomaly detection model in the following method: acquiring a first sample data set, wherein an amount of first-type data in the first sample data set is greater than that of second-type data; and using the first sample data set to train the anomaly detection model; and

training the classification model in the following method: detecting, by the anomaly detection model, an abnormal sample data set from a second sample data set; and using the abnormal sample data set to train the classification model.

5. The method according to claim 4, before using the abnormal sample data set to train the classification model, the method further comprising:

optimizing the abnormal sample data set based on a feature optimization algorithm.

6. A risk recognition method for recognizing data as secure data or risky data, wherein the method comprises:

acquiring data to be recognized, and using a preset anomaly detection model to detect whether the data to be recognized is abnormal;

if the data to be recognized is detected not to be abnormal, determining that the data to be recognized is secure data; and

if the data to be recognized is detected to be abnormal, using a preset classification model to recognize that the data to be recognized is secure data or risky data,

wherein the anomaly detection model is obtained by pre-training based on a first sample data set, and the classification model is obtained by pre-training based on a second sample data set different from the first sample data set; and

the pre-training of the classification model comprises:

detecting, by the anomaly detection model, an abnormal sample data set from the second sample data set; and

using the abnormal sample data set to train the classification model.

7. The method according to claim 6, wherein the anomaly detection model is obtained by pre-training in the following method:

acquiring the first sample data set, wherein an amount of secure data in the first sample data set is greater than that of risky data; and

using the first sample data set to train the anomaly detection model.

8. A data type recognition apparatus for recognizing data as first-type data or second-type data, wherein the apparatus comprises:

a data acquisition module for acquiring data to be recognized, and using a preset anomaly detection model to detect whether the data to be recognized is first-type data; and

a type determination module for inputting other data than the first-type data recognized by the anomaly detection model, into a classification model for recognition, wherein the classification model classifies the other data as first-type data and second-type data,

wherein the anomaly detection model is obtained by pre-training based on a first sample data set, and the classification model is obtained by pre-training based on a second sample data set different from the first sample data set; and

the apparatus further comprises: a classification model training module for detecting, by the anomaly detection model, an abnormal sample data set from the second sample data set, and using the abnormal sample data set to train the classification model.

9. The apparatus according to claim 8, further comprising an anomaly detection model training module for acquiring the first sample data set, wherein an amount of first-type data in the first sample data set is greater than that of second-type data; and using the first sample data set to train the anomaly detection model.

10. A model training apparatus for training an anomaly detection model and a classification model,

the anomaly detection model being used for detecting first-type data, and

the classification model being used for classifying other data than the first-type data detected by the anomaly detection model,

wherein the apparatus comprises:

an anomaly detection model training module for acquiring a first sample data set, wherein an amount of first-type data in the first sample data set is greater than that of second-type data; and using the first sample data set to train the anomaly detection model; and

a classification model training module for detecting, by the anomaly detection model, an abnormal sample data set from a second sample data set, and using the abnormal sample data set to train the classification model.

11. A risk recognition apparatus for recognizing data as secure data or risky data, wherein the apparatus comprises:

a data acquisition module for acquiring data to be recognized, and using a preset anomaly detection model to detect whether the data to be recognized is abnormal; and

a risk determination module for, if the data to be recognized is detected not to be abnormal, determining that the data to be recognized is secure data, and if the data to be recognized is detected to be abnormal, using a preset classification model to recognize that the data to be recognized is secure data or risky data,

wherein the anomaly detection model is obtained by pre-training based on a first sample data set, and the classification model is obtained by pre-training based on a second sample data set different from the first sample data set; and

the pre-training of the classification model comprises:

detecting, by the anomaly detection model, an abnormal sample data set from the second sample data set; and

using the abnormal sample data set to train the classification model.

12. A computer device, comprising:

a processor; and

a memory for storing a processor executable instruction,

wherein the processor is configured to:

acquire data to be recognized, and use a preset anomaly detection model to detect whether the data to be recognized is first-type data; and

input other data than the first-type data recognized by the anomaly detection model, into a classification model for recognition, wherein the classification model classifies the other data as first-type data and second-type data,

wherein the anomaly detection model is obtained by pre-training based on a first sample data set, and the classification model is obtained by pre-training based on a second sample data set different from the first sample data set; and

the pre-training of the classification model comprises:

detecting, by the anomaly detection model, an abnormal sample data set from the second sample data set; and

using the abnormal sample data set to train the classification model.

13. A computer device, comprising:

a processor; and

a memory for storing a processor executable instruction,

wherein the processor is configured to:

train an anomaly detection model in the following method: acquiring a first sample data set, wherein an amount of first-type data in the first sample data set is greater than that of second-type data; and using the first sample data set to train the anomaly detection model; and

train a classification model in the following method: detecting, by the anomaly detection model, an abnormal sample data set from a second sample data set, and using the abnormal sample data set to train the classification model.

14. A computer device, comprising:

a processor; and

a memory for storing a processor executable instruction,

wherein the processor is configured to:

acquire data to be recognized, and use a preset anomaly detection model to detect whether the data to be recognized is abnormal data;

if the data to be recognized is detected not to be abnormal, determine that the data to be recognized is secure data; and

if the data to be recognized is detected to be abnormal, use a preset classification model to

recognize that the data to be recognized is secure data or risky data,

wherein the anomaly detection model is obtained by pre-training based on a first sample data set, and the classification model is obtained by pre-training based on a second sample data set different from the first sample data set; and

the pre-training of the classification model comprises:

detecting, by the anomaly detection model, an abnormal sample data set from the second sample data set; and

using the abnormal sample data set to train the classification model.