

METHOD AND APPARATUS FOR ENCRYPTING DATA, METHOD AND APPARATUS FOR TRAINING MACHINE LEARNING MODEL, AND ELECTRONIC DEVICE

TECHNICAL FIELD

5 **[0001]** The present specification relates to the field of computer applications, and in particular, to a method and apparatus for encrypting data, a method and apparatus for training a machine learning model, and an electronic device.

BACKGROUND

10 **[0002]** With the rapid development of Internet technologies, networking and transparency of personal data of users has become an irresistible trend. Some service platforms providing the users with Internet services can collect massive user data by collecting daily service data of the users. The user data is very valuable "resources" for operators of the service platforms. The operators of the service platforms can mine a large amount of valuable information from these "resources" through data mining and machine learning. For example, in practice, with reference to a specific service scenario, data features in several dimensions can be extracted from the massive user data. The extracted features are used as training samples for training and constructing a machine learning model by using a specific machine learning algorithm. Then the trained machine learning model is used in the service scenario to guide service operations.

15

20

SUMMARY

25 **[0003]** The present specification provides a method for encrypting data, and the method includes the following: generating an $N \times M$ dimensional target matrix based on N data samples and M dimensional data features respectively corresponding to the N data samples; performing encryption calculation on the target matrix based on a Principal Component Analysis (PCA) algorithm to obtain an $N \times K$ dimensional

encryption matrix, where K is less than M ; and transmitting the encryption matrix to a modeling server, where the encryption matrix is used to train a machine learning model.

5 [0004] Optionally, the performing encryption calculation on the target matrix based on a PCA algorithm to obtain an encrypted $N \times K$ dimensional encryption matrix further includes the following: separately performing zero averaging on values of vectors corresponding to the M dimensional data features in the target matrix; calculating a covariance matrix that corresponds to the target matrix and that is obtained after the zero averaging; calculating eigenvalues of the covariance matrix and eigenvectors corresponding to the eigenvalues; sorting the calculated eigenvectors
10 based on the corresponding eigenvalues, and extracting K eigenvectors with the largest eigenvalues to generate an $M \times K$ dimensional projection matrix; and multiplying the target matrix by the projection matrix to obtain the $N \times K$ dimensional encryption matrix.

15 [0005] Optionally, the method further includes the following: locally storing the projection matrix as the encryption matrix.

[0006] Optionally, the performing encryption calculation on the target matrix based on a PCA algorithm to obtain an encrypted $N \times K$ dimensional encryption matrix includes the following: determining whether the projection matrix is locally stored; in
20 response to determining that the projection matrix is locally stored, multiplying the $N \times M$ dimensional target matrix by the projection matrix to obtain the $N \times K$ dimensional encryption matrix; and in response to determining that the projection matrix is not locally stored, performing encryption calculation on the target matrix based on the PCA algorithm to obtain the $N \times K$ dimensional encryption matrix.

25 [0007] Optionally, the method further includes the following: if the data features change in terms of a dimension, or meanings represented by the data features change, performing encryption calculation on the target matrix based on the PCA algorithm again, and updating the locally stored projection matrix based on the recalculated projection matrix.

30 [0008] The present specification further provides an apparatus for encrypting data, and the apparatus includes the following: a generation module, configured to generate an $N \times M$ dimensional target matrix based on N data samples and M dimensional data features respectively corresponding to the N data samples; a calculation module, configured to perform encryption calculation on the target matrix based on a PCA

algorithm to obtain an $N \times K$ dimensional encryption matrix, where K is less than M ; and a transmission module, configured to transmit the encryption matrix to a modeling server, where the encryption matrix is used to train a machine learning model.

5 **[0009]** Optionally, the calculation module is configured to: separately perform zero averaging on values of vectors corresponding to the M dimensional data features in the target matrix; calculate a covariance matrix that corresponds to the target matrix and that is obtained after the zero averaging; calculate eigenvalues of the covariance matrix and eigenvectors corresponding to the eigenvalues; sort the calculated
10 eigenvectors based on the corresponding eigenvalues, and extract K eigenvectors with the largest eigenvalues to generate an $M \times K$ dimensional projection matrix; and multiply the target matrix by the projection matrix to obtain the $N \times K$ dimensional encryption matrix.

[0010] Optionally, the apparatus further includes the following: a storage module,
15 configured to locally store the projection matrix as the encryption matrix.

[0011] Optionally, the calculation module is further configured to: determine whether the projection matrix is locally stored; in response to determining that the projection matrix is locally stored, multiply the $N \times M$ dimensional target matrix by the projection matrix to obtain the $N \times K$ dimensional encryption matrix; and in response
20 to determining that the projection matrix is not locally stored, perform encryption calculation on the target matrix based on the PCA algorithm to obtain the $N \times K$ dimensional encryption matrix.

[0012] Optionally, the apparatus further includes the following: an update module, configured to: if the data features change in terms of a dimension, or meanings
25 represented by the data features change, perform encryption calculation on the target matrix based on the PCA algorithm again, and update the locally stored projection matrix based on the recalculated projection matrix.

[0013] The present specification further provides a method for training a machine learning model, and the method includes the following: receiving an encryption
30 matrix transmitted by a data provider server, where the encryption matrix is an $N \times K$ dimensional encryption matrix obtained by the data provider server by performing encryption calculation on an $N \times M$ dimensional target matrix based on a PCA algorithm, and K is less than M ; and training a machine learning model by using the encryption matrix as a training sample.

[0014] Optionally, the training a machine learning model by using the encryption matrix as a training sample includes the following: combining the encryption matrix with a local training sample by using the encryption matrix as the training sample, and training the machine learning model based on a combined training sample.

5 **[0015]** The present specification further provides an apparatus for training a machine learning model, and the apparatus includes the following: a receiving module, configured to receive an encryption matrix transmitted by a data provider server, where the encryption matrix is an $N \times K$ dimensional encryption matrix obtained by the data provider server by performing encryption calculation on an $N \times M$ dimensional target matrix based on a PCA algorithm, and K is less than M ; and a training module,
10 configured to train a machine learning model by using the encryption matrix as a training sample.

[0016] Optionally, the training module is configured to: combine the encryption matrix with a local training sample by using the encryption matrix as the training sample, and train the machine learning model based on a combined training sample.
15

[0017] The present specification further provides a system for training a machine learning model, and the system includes the following: a data provider server, configured to generate an $N \times M$ dimensional target matrix based on N data samples and M dimensional data features respectively corresponding to the N data samples;
20 perform encryption calculation on the target matrix based on a PCA algorithm to obtain an $N \times K$ dimensional encryption matrix, where K is less than M ; and transmit the encryption matrix to a modeling server; and the modeling server, configured to train a machine learning model based on the encryption matrix.

[0018] Optionally, the modeling server is further configured to: combine the encryption matrix with a local training sample by using the encryption matrix as a training sample, and train the machine learning model based on a combined training sample.
25

[0019] The present specification further provides an electronic device, including the following: a processor; and a memory, configured to store a machine executable instruction.
30

[0020] By reading and executing the machine executable instruction that is stored in the memory and that corresponds to control logic for data encryption, the processor is configured to: generate an $N \times M$ dimensional target matrix based on N data samples and M dimensional data features respectively corresponding to the N data samples;

perform encryption calculation on the target matrix based on a PCA algorithm to obtain an $N \times K$ dimensional encryption matrix, where K is less than M ; and transmit the encryption matrix to a modeling server, where the encryption matrix is used to train a machine learning model.

5 **[0021]** The present specification further provides an electronic device, including the following: a processor; and a memory, configured to store a machine executable instruction.

[0022] By reading and executing the machine executable instruction that is stored in the memory and that corresponds to control logic for machine learning model
10 training, the processor is configured to: receive an encryption matrix transmitted by a data provider server, where the encryption matrix is an $N \times K$ dimensional encryption matrix obtained by the data provider server by performing encryption calculation on an $N \times M$ dimensional target matrix based on a PCA algorithm, and K is less than M ; and train a machine learning model by using the encryption matrix as a training
15 sample.

[0023] In the present specification, the $N \times M$ dimensional target matrix is generated based on the N data samples and the M dimensional data features respectively corresponding to the N data samples; encryption calculation is performed on the target matrix based on the PCA algorithm to obtain the encrypted $N \times K$
20 dimensional encryption matrix; and then the encryption matrix is transmitted to the modeling server, so that the modeling server trains the machine learning model by using the encryption matrix as the training sample.

[0024] After the encryption matrix encrypted based on the PCA algorithm is transmitted to the modeling server, the modeling server usually cannot restore an
25 original target matrix based on the encrypted encryption matrix, so that private data of a user can be protected to the greatest extent, thereby alleviating privacy disclosure of the user caused in a process of submitting a data sample to the modeling server for model training.

[0025] In addition, after encryption calculation is performed on the target matrix
30 based on the PCA algorithm, the obtained encryption matrix includes fewer dimensions than the original target matrix, so that transmission overheads for transmitting a data sample to the modeling server can be reduced. In addition, an amount of information in an original data sample can be retained to the greatest extent when encryption calculation is performed on the target matrix by using the PCA

algorithm, so that model training accuracy can still be ensured when the encrypted encryption matrix is transmitted to the modeling server for model training.

BRIEF DESCRIPTION OF DRAWINGS

[0026] FIG. 1 is a flowchart illustrating a method for encrypting data, according to an implementation of the present specification;

[0027] FIG. 2 is a schematic diagram illustrating an $N \times M$ dimensional target matrix, according to an implementation of the present specification;

[0028] FIG. 3 is a flowchart illustrating encryption calculation performed on the described target matrix based on a PCA algorithm, according to an implementation of the present specification;

[0029] FIG. 4 is a schematic diagram illustrating joint modeling performed by combining a plurality of data samples, according to an implementation of the present specification;

[0030] FIG. 5 is a flowchart illustrating a method for training a machine learning model, according to an implementation of the present specification;

[0031] FIG. 6 is a diagram illustrating a hardware structure of an electronic device including an apparatus for encrypting data, according to an implementation of the present specification;

[0032] FIG. 7 is a logical block diagram illustrating an apparatus for encrypting data, according to an implementation of the present specification;

[0033] FIG. 8 is a diagram illustrating a hardware structure of an electronic device including an apparatus for training a machine learning model, according to an implementation of the present specification; and

[0034] FIG. 9 is a logical block diagram illustrating an apparatus for training a machine learning model, according to an implementation of the present specification.

DESCRIPTION OF IMPLEMENTATIONS

[0035] In the big data era, various types of useful information can be obtained by mining massive data. Therefore, data is very important. Different institutions have respective data, but a data mining effect of any institution is subject to a data amount and a data type of the institution. A direct solution to this problem is as follows: A

plurality of institutions cooperate with each other to share data, so as to achieve a better data mining effect, thereby achieving win-win.

[0036] However, for a data owner, data is an asset of great value. In addition, for purposes such as privacy protection, the data owner is usually unwilling to provide data. In this situation, it is very difficult to perform "data sharing" in practice. Therefore, how to achieve data sharing while data security is ensured has become a top concern in the industry.

[0037] The present specification is intended to provide a technical solution. In this technical solution, original user data needed for modeling is encrypted based on a PCA algorithm to perform privacy protection on the original user data and retain an amount of information in the original user data to the greatest extent, so that user privacy can be protected without sacrificing modeling accuracy.

[0038] During implementation, M-dimensional data features can be respectively extracted from N data samples needed for modeling, and an $N \times M$ dimensional target matrix is generated based on the N data samples and the M dimensional data features respectively corresponding to the N data samples.

[0039] After the $N \times M$ dimensional target matrix is generated, encryption calculation can be performed on the target matrix based on a PCA algorithm to obtain an encrypted $N \times K$ dimensional encryption matrix, and the encryption matrix is transmitted to a modeling server as a training sample, where K is less than M.

[0040] After receiving the encryption matrix, the modeling server can train a machine learning model by using the encryption matrix as the training sample. For example, the modeling server can combine the encryption matrix with a local training sample, and then train the machine learning model based on a combined training sample.

[0041] After the encryption matrix encrypted based on the PCA algorithm is transmitted to the modeling server, the modeling server usually cannot restore an original target matrix based on the encrypted encryption matrix, so that private data of a user can be protected to the greatest extent, thereby alleviating privacy disclosure of the user caused in a process of submitting a data sample to the modeling server for model training.

[0042] In addition, after encryption calculation is performed on the target matrix based on the PCA algorithm, the obtained encryption matrix includes fewer dimensions than the original target matrix, so that transmission overheads for

transmitting a data sample to the modeling server can be reduced. In addition, an amount of information in an original data sample can be retained to the greatest extent when encryption calculation is performed on the target matrix by using the PCA algorithm, so that model training accuracy can still be ensured when the encrypted encryption matrix is transmitted to the modeling server for model training.

[0043] The following provides detailed descriptions with reference to specific application scenarios by using specific implementations.

[0044] FIG. 1 illustrates a method for encrypting data, according to an implementation of the present specification. The method is applied to a data provider server, and includes the following steps:

[0045] Step 102: Generate an $N \times M$ dimensional target matrix based on N data samples and M dimensional data features respectively corresponding to the N data samples.

[0046] Step 104: Perform dimension reduction calculation on the target matrix based on a PCA algorithm to obtain an encrypted $N \times K$ dimensional encryption matrix, where K is less than M .

[0047] Step 106: Transmit the encryption matrix to a modeling server, where the encryption matrix is used to train a machine learning model.

[0048] The data provider server can be connected to the modeling server, and provide the modeling server with a data sample needed for modeling.

[0049] For example, in practice, the data provider server and the modeling server can respectively correspond to different operators. The data provider server can transmit collected user data to the modeling server by using the collected user data as a data sample, to complete data modeling. For example, the modeling server can be a data operation platform of ALIPAY, and the data provider server can be a service platform such as a third-party bank or a courier company that is connected to the data operation platform of ALIPAY and that provides a user with Internet services.

[0050] In an initial state, the data provider server can collect daily user data at a back end, collect N pieces of user data from the collected user data as data samples, and generate an initialized data sample set based on the collected data samples.

[0051] For example, in a shown implementation, N pieces of sensitive data involving user privacy can be collected from the collected user data, and then an initialized data sample set is generated based on the sensitive data.

[0052] A specific quantity of N collected data samples is not particularly limited

in the present specification, and can be set by a person skilled in the art as needed.

[0053] A specific form of the described user data depends on a specific service scenario and a modeling demand, and is not particularly limited in the present specification. For example, in practice, if a score card model for performing risk assessment on a payment transaction initiated by the user needs to be created, the user data can be transaction data generated by the user by using a payment client.

[0054] After the data sample set is generated based on the N collected data samples, the data provider server can further preprocess the data samples in the data sample set.

[0055] Preprocessing the data samples in the data sample set usually includes the following: performing data cleaning, default value supplementation, normalization processing, or another form of preprocessing on the data samples in the data sample set. By preprocessing the data samples in the data sample set, the collected data samples can be converted into standardized data samples that are suitable for model training.

[0056] After preprocessing the data samples in the data sample set, the data provider server can respectively extract the M dimensional data features from the data samples in the data sample set.

[0057] A quantity of previously extracted M dimensional data features is not particularly limited in the present specification, and can be selected by a person skilled in the art according to actual modeling demands.

[0058] In addition, a specific type of the extracted data feature is not particularly limited in the present specification, and can be manually selected by a person skilled in the art from information actually included in the data sample according to actual modeling demands.

[0059] For example, in an implementation, the modeling server can pre-select the M dimensional data features according to actual modeling demands, and then provide the data provider server with the selected data features, so that the data provider server extracts data feature values corresponding to data features in all dimensions from the described data samples.

[0060] After separately extracting the M dimensional data features from the data samples in the data sample set, the data provider server can respectively generate data eigenvectors for the data samples based on data feature values corresponding to the extracted M dimensional data features, and then construct the N×M dimensional

target matrix based on the data eigenvectors of the data samples.

[0061] During implementation, the M dimensional data features can correspond to rows of the target matrix, or can correspond to columns of the target matrix, which is not particularly limited in the present specification.

5 [0062] Referring to FIG. 2, an example that the M dimensional data features correspond to rows of the target matrix is used. The target matrix can be represented in a form shown in FIG. 2. In the target matrix shown in FIG. 2, each column represents one data sample, and each row represents one eigenvector of the M dimensional data feature.

10 [0063] After generating the $N \times M$ dimensional target matrix based on the N data samples and the M dimensional data features respectively corresponding to the N data samples, the data provider server can perform encryption calculation on the target matrix based on the PCA algorithm to obtain the $N \times K$ dimensional encryption matrix.

[0064] The encryption matrix obtained after PCA dimension reduction calculation
15 usually cannot be restored to an original target matrix. Therefore, the encryption matrix obtained after encryption is an encrypted encryption matrix. As such, private data of a user can be protected to the greatest extent.

[0065] FIG. 3 illustrates a procedure of performing encryption calculation on the described target matrix based on a PCA algorithm, and the procedure includes the
20 following steps.

[0066] Step 302: Separately perform zero averaging on values of vectors corresponding to the M dimensional data features in the target matrix.

[0067] The zero averaging is a process of subtracting an average of a group of values from each of the group of values. In the present specification, the separately
25 performing zero averaging on values of vectors corresponding to the M dimensional data features in the target matrix is a process of separately subtracting an average of the values of the vectors from the values of the vectors corresponding to the M dimensional data features.

[0068] During implementation, the vectors corresponding to the M dimensional
30 data features in the target matrix can be successively selected as target vectors; then an average of values of the target vectors is calculated; and the average is separately subtracted from the values of the target vectors.

[0069] The target matrix shown in FIG. 2 is used as an example. An average of each row in the target matrix in FIG. 2 can be calculated, and then the average of the

row is subtracted from each value in the row.

[0070] Step 304: Calculate a covariance matrix that corresponds to the target matrix and that is obtained after the zero averaging.

[0071] The covariance matrix is a matrix consisting of covariances between
5 values of vectors in the target matrix.

[0072] When the covariance matrix that is of the target matrix and that is obtained after the zero averaging is calculated, a covariance between the vector corresponding to the M dimensional data feature and a vector corresponding to a data feature in another dimension in the target matrix can be calculated, and then the covariance
10 matrix consisting of the calculated covariance is generated based on the calculated covariance.

[0073] It is worthwhile to note that a specific process of calculating the covariance matrix is not described in detail in the present specification with reference to a specific example. A person skilled in the art can refer to related technical records
15 when implementing the technical solution described in the present specification. For example, a person skilled in the art can calculate the covariance matrix of the target matrix by using some mature tools such as matlab.

[0074] Step 306: Calculate eigenvalues of the covariance matrix and eigenvectors corresponding to the eigenvalues.

[0075] After the covariance matrix of the target matrix is calculated, the
20 eigenvalues of the covariance matrix and the eigenvectors corresponding to the eigenvalues can be further calculated. The eigenvalues of the covariance matrix usually depend on an order of the covariance matrix. For example, an $N \times M$ covariance matrix includes $N \times M$ eigenvalues.

[0076] When the eigenvalues and the eigenvectors of the covariance matrix are
25 calculated, a feature polynomial of the covariance matrix can be first obtained, and then all roots of the feature polynomial are calculated. In this case, each calculated root is an eigenvalue. After all the eigenvalues are calculated, the eigenvalues can be substituted into a system of linear equations corresponding to the covariance matrix,
30 to obtain the eigenvectors corresponding to the eigenvalues.

[0077] It is worthwhile to note that a specific process of calculating the eigenvalues of the covariance matrix and the eigenvectors corresponding to the eigenvalues is not described in detail in the present specification with reference to a specific example. A person skilled in the art can refer to related technical records

when implementing the technical solution described in the present specification. For example, a person skilled in the art can calculate the eigenvalues of the covariance matrix and the eigenvectors corresponding to the eigenvalues by using some mature tools such as matlab.

5 **[0078]** Step 308: Sort the calculated eigenvectors based on the corresponding eigenvalues, and extract K eigenvectors with the largest eigenvalues to generate an $M \times K$ dimensional projection matrix.

[0079] After all the eigenvalues of the covariance matrix and the eigenvectors corresponding to the eigenvalues are calculated, the eigenvectors can be further sorted
10 based on the corresponding eigenvalues, for example, sorted in descending order. After the eigenvectors are sorted based on the eigenvalues, the first K eigenvectors with the largest eigenvalues can be extracted to generate the $M \times K$ dimensional projection matrix.

[0080] Here, K can be less than M. In practice, K can be manually specified by a
15 person skilled in the art with reference to actual demands.

[0081] Step 310: Multiply the target matrix by the projection matrix to obtain the encrypted $N \times K$ dimensional encryption matrix.

[0082] The $M \times K$ dimensional projection matrix is a projection matrix obtained by finally encrypting an original target matrix. When the original $N \times M$ target matrix is
20 encrypted based on the $M \times K$ dimensional projection matrix, the original high-dimensional target matrix can be mapped to low-dimensional projection matrix space.

[0083] During implementation, the original $N \times M$ target matrix can be mapped to the $M \times K$ dimensional projection matrix space by multiplying the original $N \times M$ target
25 matrix by the $M \times K$ dimensional projection matrix (in other words, through linear projection). The original $N \times M$ target matrix can be right-multiplied or left-multiplied by the $M \times K$ dimensional projection matrix.

[0084] For example, if the M dimensional data features are used as columns of the target matrix, the original $N \times M$ target matrix can be right-multiplied by the $M \times K$
30 dimensional projection matrix, and the original $N \times M$ target matrix can be mapped to the $M \times K$ dimensional projection matrix space. Therefore, during implementation, the original $N \times M$ target matrix can be left-multiplied by the $M \times K$ dimensional projection matrix, a left multiplication result is transposed, and the original $N \times M$ target matrix is mapped to the $M \times K$ dimensional projection matrix space.

[0085] After the original $N \times M$ target matrix is mapped to the $M \times K$ dimensional projection matrix space, the $N \times K$ dimensional encryption matrix can be obtained. In this case, the encryption matrix is a data sample obtained after the $M \times K$ dimensional projection matrix is encrypted.

5 **[0086]** In a shown implementation, the data provider server can further locally store the projection matrix as the encryption matrix after calculating the $M \times K$ dimensional projection matrix by using the previously shown calculation process.

[0087] Subsequently, the data provider server collects the latest N data samples again; generates an $N \times M$ dimensional matrix based on the N data samples and M dimensional data features respectively corresponding to the N data samples; and then
10 can determine whether the projection matrix is locally stored.

[0088] In response to determining that the projection matrix is locally stored, the data provider server encrypts the $N \times M$ matrix by directly using the stored projection matrix. A specific encryption process is not described in detail again.

15 **[0089]** Certainly, in response to determining that the projection matrix is not locally stored, the data provider server can perform, based on the described PCA algorithm based dimension reduction process, encryption calculation on the described encryption matrix to generate the described projection matrix.

[0090] In addition, it is worthwhile to note that, in practice, if the M dimensional data features change in terms of a dimension (for example, a data feature in a new dimension is added or a data feature in some dimensions is deleted), or meanings represented by some or all of the M dimensional data features change, the data provider server can recalculate the projection matrix based on the described PCA algorithm based encryption calculation process, and update the locally stored
20 projection matrix by using the recalculated projection matrix.

[0091] As such, a locally stored invalid encryption matrix can be updated in time when a data feature needed for modeling is updated, thereby preventing modeling accuracy from being affected by a loss of an amount of data information when the original target matrix is encrypted by using the invalid encryption matrix.

30 **[0092]** In the present specification, after dimension reduction calculation is performed on the target matrix based on the described PCA algorithm based encryption calculation process to obtain the encrypted $N \times K$ dimensional encryption matrix, the data provider server can transmit the encryption matrix as a training sample to the modeling server connected to the data provider server.

[0093] After receiving the encryption matrix transmitted by the data provider server, the modeling server can train a machine learning model by using the encryption matrix as a training sample.

[0094] In a shown implementation, the modeling server can combine the encryption matrix with a locally stored training sample, and then train the machine learning model based on a combined training sample.

[0095] FIG. 4 is a schematic diagram illustrating joint modeling performed by combining a plurality of data samples, according to the present specification.

[0096] In a scenario, the described modeling server can be a data operation platform of ALIPAY, and the data provider server can include a service platform such as a bank or a third-party financial institution that is connected to the data operation platform of ALIPAY and that provides a user with Internet services. In practice, the data operation platform of ALIPAY is a non-credited third party for the data provider server. Therefore, the data provider server directly provides local user transaction data for the data operation platform of ALIPAY for data modeling. Consequently, user privacy may be disclosed during data transmission. In this case, each data provider server can perform, based on a PCA algorithm and the projection matrix, encryption calculation on an $N \times M$ dimensional target matrix generated based on an original transaction data sample, to obtain an $N \times K$ dimensional encryption matrix, and then transmit the encryption matrix to the data operation platform of ALIPAY as a training sample. The data operation platform of ALIPAY can combine a received training sample provided by the data provider server with a localized data sample, and then train the machine learning model based on a combined training sample, for example, combine user transaction data provided by a bank or a third-party financial institution with localized user transaction data in the data operation platform of ALIPAY to jointly train a score card model for assessing a transaction risk of the user.

[0097] It is worthwhile to note that a specific type of the machine learning model is not particularly limited in the present specification. For example, in practice, the machine learning model can be a supervised prediction model established based on a supervised machine learning algorithm (such as a regression algorithm), for example, a score card model that is used to assess a transaction risk of the user and that is trained based on payment transaction data of the user; or can be an unsupervised classification model established based on an unsupervised machine learning algorithm (for example, a k-means algorithm), for example, a recommendation model that is

used to deliver an advertisement or page content to the user and that is trained based on a click and access data of the user.

[0098] After the machine learning model is trained by using the previously described modeling methods, the data provider server can still subsequently encrypt, based on the projection matrix, a data matrix constructed based on a collected data sample and a related data feature, and then transmit the data matrix to the machine learning model for calculation, to obtain an output result of the model. For example, the machine learning model is the score card model that is used to assess the transaction risk of the user and that is trained based on the payment transaction data of the user. Based on the projection matrix, the data provider server can perform encryption calculation on the data matrix constructed based on the collected user transaction data, and then transmit the data matrix to the score card model as input data, to obtain a risk score corresponding to each transaction.

[0099] The above describes the method for encrypting data provided in the present implementation of the present specification. Referring to FIG. 5, based on a same idea, an implementation of the present specification provides a method for training a machine learning model. The method is applied to a modeling server, and includes the following steps:

[0100] Step 502: Receive an encryption matrix transmitted by a data provider server, where the encryption matrix is an $N \times K$ dimensional encryption matrix obtained by the data provider server by performing encryption calculation on an $N \times M$ dimensional target matrix based on a PCA algorithm, and K is less than M .

[0101] Step 504: Train a machine learning model by using the encryption matrix as a training sample.

[0102] For an implementation process corresponding to the technical features in the steps shown in FIG. 5, details are omitted in the present implementation, and references can be made to the descriptions in the previous implementation.

[0103] It can be seen from the described implementations that the $N \times M$ dimensional target matrix is generated based on the N data samples and the M dimensional data features respectively corresponding to the N data samples; encryption calculation is performed on the target matrix based on the PCA algorithm to obtain the $N \times K$ dimensional encryption matrix; and then the encryption matrix is transmitted to the modeling server, so that the modeling server trains the machine learning model by using the encryption matrix as the training sample.

[0104] After the encryption matrix encrypted based on the PCA algorithm is transmitted to the modeling server, the modeling server usually cannot restore an original target matrix based on the encrypted encryption matrix, so that private data of a user can be protected to the greatest extent, thereby alleviating privacy disclosure of the user caused in a process of submitting a data sample to the modeling server for model training.

[0105] In addition, after encryption calculation is performed on the target matrix based on the PCA algorithm, the obtained encryption matrix includes fewer dimensions than the original target matrix, so that transmission overheads for transmitting a data sample to the modeling server can be reduced.

[0106] In addition, an amount of information in an original data sample can be retained to the greatest extent when encryption calculation is performed on the target matrix by using the PCA algorithm, so that model training accuracy can still be ensured when the encrypted encryption matrix is transmitted to the modeling server for model training.

[0107] The present specification further provides an implementation of an apparatus for encrypting data corresponding to the described method implementation.

[0108] The implementation of the apparatus for encrypting data in the present specification can be applied to an electronic device. The apparatus implementation can be implemented by software, hardware, or a combination of hardware and software. Software implementation is used as an example. As a logic apparatus, the apparatus is formed by reading a corresponding computer program instruction in a nonvolatile memory to a memory by a processor of the electronic device that the apparatus is located in. In terms of hardware, FIG. 6 is a hardware structural diagram illustrating an electronic device that an apparatus for encrypting data is located in, according to the present specification. In addition to a processor, a memory, a network interface, and a nonvolatile memory shown in FIG. 6, the electronic device that the apparatus in the present implementation is located in usually can include other hardware based on an actual function of the electronic device. Details are omitted here for simplicity.

[0109] FIG. 7 is a block diagram illustrating an apparatus for encrypting data, according to an example implementation of the present specification.

[0110] Referring to FIG. 7, the apparatus 70 for encrypting data can be applied to the electronic device shown in FIG. 6, and include a generation module 701, a

calculation module 702, and a transmission module 703.

[0111] The generation module 701 is configured to generate an $N \times M$ dimensional target matrix based on N data samples and M dimensional data features respectively corresponding to the N data samples.

5 [0112] The calculation module 702 is configured to perform encryption calculation on the target matrix based on a PCA algorithm to obtain an $N \times K$ dimensional encryption matrix, where K is less than M .

[0113] The transmission module 703 is configured to transmit the encryption matrix to a modeling server, where the encryption matrix is used to train a machine
10 learning model.

[0114] In the present implementation, the calculation module 702 is configured to: separately perform zero averaging on values of vectors corresponding to the M dimensional data features in the target matrix; calculate a covariance matrix that corresponds to the target matrix and that is obtained after the zero averaging; calculate
15 eigenvalues of the covariance matrix and eigenvectors corresponding to the eigenvalues; sort the calculated eigenvectors based on the corresponding eigenvalues, and extract K eigenvectors with the largest eigenvalues to generate an $M \times K$ dimensional projection matrix; and multiply the target matrix by the projection matrix to obtain the $N \times K$ dimensional encryption matrix.

20 [0115] In the present implementation, the apparatus further includes the following: a storage module 704 (not shown in FIG. 7), configured to locally store the projection matrix as the encryption matrix.

[0116] In the present implementation, the calculation module 702 is further configured to: determine whether the projection matrix is locally stored; in response
25 to determining that the projection matrix is locally stored, multiply the $N \times M$ dimensional target matrix by the projection matrix to obtain the $N \times K$ dimensional encryption matrix; and in response to determining that the projection matrix is not locally stored, perform encryption calculation on the target matrix based on the PCA algorithm to obtain the $N \times K$ dimensional encryption matrix.

30 [0117] In the present implementation, the apparatus 70 further includes the following: an update module 705 (not shown in FIG. 7), configured to: if the data features change in terms of a dimension, or meanings represented by the data features change, perform encryption calculation on the target matrix based on the PCA algorithm again, and update the locally stored projection matrix based on the

recalculated projection matrix.

[0118] The present specification further provides an implementation of an apparatus for training a machine learning model corresponding to the described method implementation.

5 **[0119]** The implementation of the apparatus for training a machine learning model in the present specification can be applied to an electronic device. The apparatus implementation can be implemented by software, hardware, or a combination of hardware and software. Software implementation is used as an example. As a logic apparatus, the apparatus is formed by reading a corresponding computer program
10 instruction in a nonvolatile memory to a memory by a processor of the electronic device that the apparatus is located in. In terms of hardware, FIG. 8 is a hardware structural diagram illustrating an electronic device that an apparatus for training a machine learning model is located in, according to the present specification. In addition to a processor, a memory, a network interface, and a nonvolatile memory
15 shown in FIG. 8, the electronic device that the apparatus in the present implementation is located in usually can include other hardware based on an actual function of the electronic device. Details are omitted here for simplicity.

[0120] FIG. 9 is a block diagram illustrating an apparatus for training a machine learning model, according to an example implementation of the present specification.

20 **[0121]** Referring to FIG. 9, the apparatus 90 for training a machine learning model can be applied to the electronic device shown in FIG. 8, and include a receiving module 901 and a training module 902.

[0122] The receiving module 901 is configured to receive an encryption matrix transmitted by a data provider server, where the encryption matrix is an $N \times K$
25 dimensional encryption matrix obtained by the data provider server by performing encryption calculation on an $N \times M$ dimensional target matrix based on a PCA algorithm, and K is less than M .

[0123] The training module 902 is configured to train a machine learning model by using the encryption matrix as a training sample.

30 **[0124]** In the present implementation, the training module 902 is further configured to: combine the encryption matrix with a local training sample by using the encryption matrix as the training sample, and train the machine learning model based on a combined training sample.

[0125] For implementation processes of functions and roles of the modules in the

apparatus, references can be made to implementation processes of corresponding steps in the described method. Details are omitted here for simplicity.

[0126] Because an apparatus implementation basically corresponds to a method implementation, for related parts, references can be made to partial descriptions in the method implementation. The described apparatus implementation is merely an example. The units described as separate parts can or cannot be physically separate, and parts displayed as units can or cannot be physical units, in other words, can be located in one position, or can be distributed in a plurality of network units. Some or all of the modules can be selected based on actual demands to achieve the objectives of the solutions of the present specification. A person of ordinary skill in the art can understand and implement the implementations of the present specification without creative efforts.

[0127] The system, apparatus, module, or unit illustrated in the previous implementations can be implemented by using a computer chip or an entity, or can be implemented by using a product having a certain function. A typical implementation device is a computer, and the computer can be a personal computer, a laptop computer, a cellular phone, a camera phone, a smartphone, a personal digital assistant, a media player, a navigation device, an email receiving and sending device, a game console, a tablet computer, a wearable device, or any combination of these devices.

[0128] The present specification further provides an implementation of a system for training a machine learning model corresponding to the described method implementation.

[0129] The system for training a machine learning model can include a data provider server and a modeling server.

[0130] The data provider server can generate an $N \times M$ dimensional target matrix based on N data samples and M dimensional data features respectively corresponding to the N data samples; perform encryption calculation on the target matrix based on a PCA algorithm to obtain an $N \times K$ dimensional encryption matrix, where K is less than M ; and transmit the encryption matrix to the modeling server.

[0131] The modeling server is configured to train a machine learning model based on the encryption matrix.

[0132] In the present implementation, the modeling server is further configured to: combine the encryption matrix with a local training sample by using the encryption matrix as a training sample, and train the machine learning model based on a

combined training sample.

[0133] The present specification further provides an implementation of an electronic device corresponding to the described method implementation. The electronic device includes a processor and a memory configured to store a machine
5 executable instruction. The processor and the memory are generally connected to each other by using an internal bus. In other possible implementations, the device can further include an external interface that is capable of communicating with another device or component.

[0134] In the present implementation, by reading and executing the machine
10 executable instruction that is stored in the memory and that corresponds to control logic for data encryption shown in FIG. 1, the processor is configured to: generate an $N \times M$ dimensional target matrix based on N data samples and M dimensional data features respectively corresponding to the N data samples; perform encryption calculation on the target matrix based on a PCA algorithm to obtain an $N \times K$
15 dimensional encryption matrix, where K is less than M ; and transmit the encryption matrix to a modeling server, where the encryption matrix is used to train a machine learning model.

[0135] In the present example, by reading and executing the machine executable instruction that is stored in the memory and that corresponds to control logic for data
20 encryption shown in FIG. 1, the processor is further configured to: separately perform zero averaging on values of vectors corresponding to the M dimensional data features in the target matrix; calculate a covariance matrix that corresponds to the target matrix and that is obtained after the zero averaging; calculate eigenvalues of the covariance matrix and eigenvectors corresponding to the eigenvalues; sort the calculated
25 eigenvectors based on the corresponding eigenvalues, and extract K eigenvectors with the largest eigenvalues to generate an $M \times K$ dimensional projection matrix; and multiply the target matrix by the projection matrix to obtain the encrypted $N \times K$ dimensional encryption matrix.

[0136] In the present example, by reading and executing the machine executable
30 instruction that is stored in the memory and that corresponds to control logic for data encryption shown in FIG. 1, the processor is further configured to: locally store the projection matrix as the encryption matrix.

[0137] In the present example, by reading and executing the machine executable instruction that is stored in the memory and that corresponds to control logic for data

encryption shown in FIG. 1, the processor is further configured to: determine whether the projection matrix is locally stored; in response to determining that the projection matrix is locally stored, multiply the $N \times M$ dimensional target matrix by the projection matrix to obtain the $N \times K$ dimensional encryption matrix; and in response to
 5 determining that the projection matrix is not locally stored, perform encryption calculation on the target matrix based on the PCA algorithm to obtain the $N \times K$ dimensional encryption matrix.

[0138] In the present example, by reading and executing the machine executable instruction that is stored in the memory and that corresponds to control logic for data
 10 encryption, the processor is further configured to: if the data features change in terms of a dimension, or meanings represented by the data features change, perform encryption calculation on the target matrix based on the PCA algorithm again, and update the locally stored projection matrix based on the recalculated projection matrix.

[0139] The present specification further provides an implementation of another electronic device corresponding to the described method implementation. The electronic device includes a processor and a memory configured to store a machine executable instruction. The processor and the memory are generally connected to each other by using an internal bus. In other possible implementations, the device can
 20 further include an external interface that is capable of communicating with another device or component.

[0140] In the present implementation, by reading and executing the machine executable instruction that is stored in the memory and that corresponds to control logic for machine learning model training shown in FIG. 5, the processor is
 25 configured to: receive an encryption matrix transmitted by a data provider server, where the encryption matrix is an $N \times K$ dimensional encryption matrix obtained by the data provider server by performing encryption calculation on an $N \times M$ dimensional target matrix based on a PCA algorithm, and K is less than M ; and train a machine learning model by using the encryption matrix as a training sample.

[0141] In the present implementation, by reading and executing the machine executable instruction that is stored in the memory and that corresponds to control logic for machine learning model training shown in FIG. 5, the processor is further
 30 configured to: combine the encryption matrix with a local training sample by using the encryption matrix as the training sample, and train the machine learning model

based on a combined training sample.

[0142] A person skilled in the art can easily figure out another implementation of the present specification after considering the present specification and practicing the present disclosure. The present specification is intended to cover any variations, functions, or adaptive changes of the present specification. These variations, functions, or adaptive changes comply with general principles of the present specification, and include common knowledge or a commonly used technical means in the technical field that is not disclosed in the present specification. The specification and the implementations are merely considered as examples, and the actual scope and the spirit of the present specification are described by the following claims.

[0143] It should be understood that the present specification is not limited to the precise structures that have been described above and shown in the accompanying drawings, and various modifications and changes can be made without departing from the scope of the present specification. The scope of the present specification is limited only by the appended claims.

[0144] Specific implementations of the present specification are described above. Other implementations fall within the scope of the appended claims. In some situations, the actions or steps described in the claims can be performed in an order different from the order in the implementations and the desired results can still be achieved. In addition, the process depicted in the accompanying drawings does not necessarily need a particular execution order to achieve the desired results. In some implementations, multitasking and parallel processing can be advantageous.

[0145] The previous descriptions are merely examples of implementations of the present application, but are not intended to limit the present application. Any modification, equivalent replacement, or improvement made without departing from the spirit and principle of the present application should fall within the protection scope of the present application.

CLAIMS

What is claimed is:

1. A method for encrypting data, wherein the method comprises:
generating an $N \times M$ dimensional target matrix based on N data samples and M
5 dimensional data features respectively corresponding to the N data samples;
performing encryption calculation on the target matrix based on a PCA algorithm
to obtain an $N \times K$ dimensional encryption matrix, wherein K is less than M ; and
transmitting the encryption matrix to a modeling server, wherein the encryption
matrix is used to train a machine learning model.
- 10 2. The method according to claim 1, wherein the performing encryption
calculation on the target matrix based on a PCA algorithm to obtain an encrypted
 $N \times K$ dimensional encryption matrix further comprises:
separately performing zero averaging on values of vectors corresponding to the
 M dimensional data features in the target matrix;
15 calculating a covariance matrix that corresponds to the target matrix and that is
obtained after the zero averaging;
calculating eigenvalues of the covariance matrix and eigenvectors corresponding
to the eigenvalues;
sorting the calculated eigenvectors based on the corresponding eigenvalues, and
20 extracting K eigenvectors with the largest eigenvalues to generate an $M \times K$
dimensional projection matrix; and
multiplying the target matrix by the projection matrix to obtain the encrypted
 $N \times K$ dimensional encryption matrix.
- 25 3. The method according to claim 2, further comprising:
locally storing the projection matrix as the encryption matrix.
4. The method according to claim 3, wherein the performing encryption
calculation on the target matrix based on a PCA algorithm to obtain an encrypted
 $N \times K$ dimensional encryption matrix comprises:
determining whether the projection matrix is locally stored;
30 in response to determining that the projection matrix is locally stored,
multiplying the $N \times M$ dimensional target matrix by the projection matrix to obtain the
 $N \times K$ dimensional encryption matrix; and

in response to determining that the projection matrix is not locally stored, performing encryption calculation on the target matrix based on the PCA algorithm to obtain the $N \times K$ dimensional encryption matrix.

5. The method according to claim 3, further comprising:

5 if the data features change in terms of a dimension, or meanings represented by the data features change, performing encryption calculation on the target matrix based on the PCA algorithm again, and updating the locally stored projection matrix based on the recalculated projection matrix.

6. An apparatus for encrypting data, wherein the apparatus comprises:

10 a generation module, configured to generate an $N \times M$ dimensional target matrix based on N data samples and M dimensional data features respectively corresponding to the N data samples;

a calculation module, configured to perform encryption calculation on the target matrix based on a PCA algorithm to obtain an $N \times K$ dimensional encryption matrix, 15 wherein K is less than M ; and

a transmission module, configured to transmit the encryption matrix to a modeling server, wherein the encryption matrix is used to train a machine learning model.

7. The apparatus according to claim 6, wherein the calculation module is 20 configured to:

separately perform zero averaging on values of vectors corresponding to the M dimensional data features in the target matrix;

calculate a covariance matrix that corresponds to the target matrix and that is obtained after the zero averaging;

25 calculate eigenvalues of the covariance matrix and eigenvectors corresponding to the eigenvalues;

sort the calculated eigenvectors based on the corresponding eigenvalues, and extract K eigenvectors with the largest eigenvalues to generate an $M \times K$ dimensional projection matrix; and

30 multiply the target matrix by the projection matrix to obtain the $N \times K$ dimensional encryption matrix.

8. The apparatus according to claim 7, further comprising:

a storage module, configured to locally store the projection matrix as the encryption matrix.

9. The apparatus according to claim 8, wherein the calculation module is further configured to:

determine whether the projection matrix is locally stored;

in response to determining that the projection matrix is locally stored, multiply
5 the $N \times M$ dimensional target matrix by the projection matrix to obtain the $N \times K$ dimensional encryption matrix; and

in response to determining that the projection matrix is not locally stored, perform encryption calculation on the target matrix based on the PCA algorithm to obtain the $N \times K$ dimensional encryption matrix.

10 10. The apparatus according to claim 8, further comprising:

an update module, configured to: if the data features change in terms of a dimension, or meanings represented by the data features change, perform encryption calculation on the target matrix based on the PCA algorithm again, and update the locally stored projection matrix based on the recalculated projection matrix.

15 11. A method for training a machine learning model, wherein the method comprises:

receiving an encryption matrix transmitted by a data provider server, wherein the encryption matrix is an $N \times K$ dimensional encryption matrix obtained by the data provider server by performing encryption calculation on an $N \times M$ dimensional target
20 matrix based on a PCA algorithm, and K is less than M ; and

training a machine learning model by using the encryption matrix as a training sample.

12. The method according to claim 11, wherein the training a machine learning model by using the encryption matrix as a training sample comprises:

25 combining the encryption matrix with a local training sample by using the encryption matrix as the training sample, and training the machine learning model based on a combined training sample.

13. An apparatus for training a machine learning model, wherein the apparatus comprises:

30 a receiving module, configured to receive an encryption matrix transmitted by a data provider server, wherein the encryption matrix is an $N \times K$ dimensional encryption matrix obtained by the data provider server by performing encryption calculation on an $N \times M$ dimensional target matrix based on a PCA algorithm, and K is less than M ; and

a training module, configured to train a machine learning model by using the encryption matrix as a training sample.

14. The apparatus according to claim 13, wherein the training module is configured to:

5 combine the encryption matrix with a local training sample by using the encryption matrix as the training sample, and train the machine learning model based on a combined training sample.

15. A system for training a machine learning model, wherein the system comprises:

10 a data provider server, configured to generate an $N \times M$ dimensional target matrix based on N data samples and M dimensional data features respectively corresponding to the N data samples; perform encryption calculation on the target matrix based on a PCA algorithm to obtain an $N \times K$ dimensional encryption matrix, wherein K is less than M ; and transmit the encryption matrix to a modeling server; and

15 the modeling server, configured to train a machine learning model based on the encryption matrix.

16. The system according to claim 15, wherein the modeling server is further configured to:

20 combine the encryption matrix with a local training sample by using the encryption matrix as a training sample, and train the machine learning model based on a combined training sample.

17. An electronic device, comprising:

a processor; and

a memory, configured to store a machine executable instruction, wherein

25 by reading and executing the machine executable instruction that is stored in the memory and that corresponds to control logic for data encryption, the processor is configured to:

generate an $N \times M$ dimensional target matrix based on N data samples and M dimensional data features respectively corresponding to the N data samples;

30 perform encryption calculation on the target matrix based on a PCA algorithm to obtain an $N \times K$ dimensional encryption matrix, wherein K is less than M ; and

transmit the encryption matrix to a modeling server, wherein the encryption matrix is used to train a machine learning model.

18. An electronic device, comprising:
a processor; and
a memory, configured to store a machine executable instruction, wherein
by reading and executing the machine executable instruction that is stored in the
5 memory and that corresponds to control logic for machine learning model training, the
processor is configured to:
- receive an encryption matrix transmitted by a data provider server, wherein
the encryption matrix is an $N \times K$ dimensional encryption matrix obtained by the data
provider server by performing encryption calculation on an $N \times M$ dimensional target
10 matrix based on a PCA algorithm, and K is less than M ; and
train a machine learning model by using the encryption matrix as a training
sample.