CLAIMS

What is claimed is:

1. A trusted execution environment (TEE)-based model training method, comprising:

obtaining multiple encrypted target samples from an encrypted training sample set;

inputting the multiple encrypted target samples into a TEE, decrypting the encrypted target samples in the TEE, and inputting the decrypted target samples into a feature extraction model to obtain sample features;

determining the sample features output in the TEE as target sample features used for a current iteration; and

performing current iterative training on a target model by using the target sample features, wherein the performing current iterative training on a target model by using the target sample features comprises: determining a first gradient value corresponding to the target model in the current iteration based on a loss function of the target model; updating model parameters of the target model by using the first gradient value; calculating a second gradient value corresponding to the feature extraction model in the current iteration based on the first gradient value; and inputting the second gradient value into the TEE, and updating model parameters of the feature extraction model in the TEE by using the second gradient value; and wherein the target model is iteratively trained based on a plurality of training processes; and a procedure of obtaining multiple encrypted target samples by each training process comprises: during a first iteration, obtaining a training sample subset corresponding to a current training process, and determining encrypted training samples in the training sample subset as encrypted target samples of the current iteration; and during a non-first iteration, exchanging some encrypted training samples with a designated training process to obtain encrypted target samples of the current iteration.

2. The method according to claim 1, wherein, before the inputting the decrypted target samples into a feature extraction model, the method further comprises:

performing feature enhancement processing on the decrypted target samples in the TEE.

3. The method according to claim 1, wherein the exchanging some encrypted training samples with a designated training process comprises:

sending, by each training process, encrypted training samples with a predetermined proportion in a corresponding training sample subset to a first designated training process; and

receiving encrypted training samples sent by a second designated training process.

4. The method according to claim 1, wherein each training process is used to obtain designated model parameters of the target model through training.

5. The method according to claim 1, wherein after the performing current iterative training on a target model by using the target sample features, the method further comprises:

performing the following operation for each training process:

exchanging some target sample features with a designated training process to obtain exchanged target sample features of a current training process for a next iteration, and training the target model by using the exchanged target sample features in the next iteration.

6. A TEE-based model training apparatus, comprising:

an acquisition unit, configured to obtain multiple encrypted target samples from an encrypted training sample set;

an input unit, configured to input the multiple encrypted target samples into a TEE, decrypt the encrypted target samples in the TEE, and input the decrypted target samples into a feature extraction model to obtain sample features;

a determining unit, configured to determine the sample features output in the TEE as target sample features used for a current iteration; and

a training unit, configured to perform current iterative training on a target model by using the target sample features, wherein the training unit is configured to: determine a first gradient value corresponding to the target model in the current iteration based on a loss function of the target model; and update model parameters of the target model by using the first gradient value; and the training unit is further configured to: calculate a

second gradient value corresponding to the feature extraction model in the current iteration based on the first gradient value; and update model parameters of the feature extraction model by using the second gradient value; and wherein the target model is iteratively trained based on a plurality of training processes; and when the acquisition unit obtains the multiple encrypted target sample, each training process performs the following operations: during a first iteration, obtaining a training sample subset corresponding to a current training process, and determining encrypted training samples in the training sample subset as encrypted target samples of the current iteration; and during a non-first iteration, exchanging some encrypted training samples with a designated training process to obtain encrypted target samples of the current iteration.

7. The apparatus according to claim 6, wherein before inputting the decrypted target samples into the feature extraction model, the input unit is further configured to:

perform feature enhancement processing on the decrypted target samples in the TEE.

8. The apparatus according to claim 6, wherein when the acquisition unit exchanges some encrypted training samples with the designated training process, each training process sends encrypted training samples with a predetermined proportion in a corresponding training sample subset to a first designated training process; and

receives encrypted training samples sent by a second designated training process.

9. The apparatus according to claim 6, wherein each training process is used to obtain designated model parameters of the target model through training.

10. The apparatus according to claim 6, wherein the training unit is further configured to:

perform the following operation for each training process:

exchanging some target sample features with a designated training process to obtain exchanged target sample features of a current training process for a next iteration, and training the target model by using the exchanged target sample features in the next iteration.

11. A TEE-based model training apparatus, comprising:

a processor; and

a memory configured to store machine executable instructions, wherein

by reading and executing the machine executable instructions that are stored in the memory and that correspond to TEE-based model training logic, the processor is enabled to:

obtain multiple encrypted target samples from an encrypted training sample set;

input the multiple encrypted target samples into a TEE, decrypt the encrypted target samples in the TEE, and input the decrypted target samples into a feature extraction model to obtain sample features;

determine the sample features output in the TEE as target sample features used for a current iteration; and

perform current iterative training on a target model by using the target sample features, wherein when performing current iterative training on the target model by using the target sample features, the processor is enabled to determine a first gradient value corresponding to the target model in the current iteration based on a loss function of the target model, and update model parameters of the target model by using the first gradient value; wherein the processor is further enabled to calculate a second gradient value corresponding to the feature extraction model in the current iteration based on the first gradient value, and update model parameters of the feature extraction model in the TEE by using the second gradient value; and wherein the target model is iteratively trained based on a plurality of training processes; and when each training process obtains multiple target samples used for the current iteration, the processor is enabled to, during a first iteration, obtain a training sample subset corresponding to a current training process, and determine encrypted training samples in the training sample subset as encrypted target samples of the current iteration, and during a non-first iteration, exchange some encrypted training samples with a designated training process to obtain encrypted target samples of the current iteration.