

CLAIMS

What is claimed is:

1. A computer-implemented method for trainings of learning models, comprising:
receiving (102) current streaming sample data; and
5 training (104) a current deep learning model based on the current streaming sample data, the training comprising:

10 before training the current deep learning model, training (204A) a shallow learning model that is a separate machine learning model obtained through online training based on historical sample data associated with the current streaming sample data, wherein:

the historical sample data is historical streaming sample data before the current streaming sample data, and

a structure of the shallow learning model is simpler than the structure of the current deep learning model; and

15 after the training of the shallow learning model is completed, training (204B) the current deep learning model based on the current streaming sample data,
wherein parameters of ~~a~~-the shallow learning model are used as initialization parameters of the current deep learning model, ~~and the shallow learning model is obtained through training based on historical sample data associated with the current streaming sample data.~~
20 current streaming sample data.

~~2. The method according to claim 1, wherein the historical sample data is historical streaming sample data before the current streaming sample data, and the shallow learning model is obtained through online training based on the historical streaming sample data.~~

25 ~~32.~~ The method according to claim 1, wherein the historical sample data is offline sample data, and the shallow learning model is obtained through offline training based on the offline sample data.

~~43.~~ The method according to ~~any one of claims 1 to 3~~claim 1 or 2, further comprising:

30 after the training of the current deep learning model is finished to obtain a trained deep learning model:

in response to determining that performance of the trained deep learning model is

improved compared to performance of the current deep learning model, using (208B) the trained deep learning model as the latest deep learning model; or

in response to determining that performance of the trained deep learning model is not improved compared to performance of the current deep learning model,
5 performing the following operations:

increasing (210B) ~~the~~ a quantity of hidden layers of the current deep learning model to obtain a deep learning model with the increased quantity of hidden layers;

training (212B) the deep learning model with the increased quantity of
10 hidden layers based on the current streaming sample data, to obtain a new deep learning model; and

determining the latest deep learning model based on a result of performance comparison between the new deep learning model and the current deep learning model.

15 ~~54~~. The method according to claim ~~43~~, wherein the parameters of the shallow learning model are used as initialization parameters of the deep learning model with the increased quantity of hidden layers.

~~65~~. The method according to claim ~~43~~ or ~~54~~, wherein the determining the latest deep learning model based on a result of performance comparison between the new
20 deep learning model and the trained deep learning model comprises:

in response to determining that performance of the new deep learning model is improved compared to the performance of the current deep learning model, using (216B) the new deep learning model as the latest deep learning model; or

in response to determining that performance of the new deep learning model
25 is not improved compared to the performance of the current deep learning model, using (218B) the current deep learning model as the latest deep learning model.

~~76~~. The method according to any one of claims ~~43~~ to ~~65~~, further comprising:

weighting the latest deep learning model and the shallow learning model to obtain the latest learning model.

30 ~~87~~. An apparatus (300) for trainings of learning models, comprising:

a receiving unit (302), configured to receive (102) current streaming sample data;
and

a training unit (304), configured to train (104) a current deep learning model based on the current streaming sample data, the training comprising:

before training the current deep learning model, training (204A) a shallow learning model that is a separate machine learning model obtained through online training based on historical sample data associated with the current streaming sample data, wherein:

5 the historical sample data is historical streaming sample data before the current streaming sample data, and

a structure of the shallow learning model is simpler than the structure of the current deep learning model; and

10 after the training of the shallow learning model is completed, training the current deep learning model based on the current streaming sample data, wherein parameters of a-the shallow learning model are used as initialization parameters of the current deep learning model, and the shallow learning model is obtained through training based on historical sample data associated with the current streaming sample data.

15 ~~9. The apparatus according to claim 8, wherein the historical sample data is historical streaming sample data before the current streaming sample data, and the shallow learning model is obtained through online training based on the historical streaming sample data.~~

20 ~~108.~~ The apparatus (300) according to claim ~~8~~7, wherein the historical sample data is offline sample data, and the shallow learning model is obtained through offline training based on the offline sample data.

~~119.~~ The apparatus (300) according to ~~any one of claims 8 to 10~~claim 7 or 8, further comprising an evaluation unit (306); and

25 after the training unit (304) finishes the training of the current deep learning model to obtain a trained deep learning model:

in response to determining that performance of the trained deep learning model is improved compared to performance of the current deep learning model, the evaluation unit (306) is configured to use (208B) the trained deep learning model as the latest deep learning model; or

30 in response to determining that performance of the trained deep learning model is not improved compared to performance of the current deep learning model:

the training unit (304) is further configured to increase (210B) ~~the a~~ quantity of hidden layers of the current deep learning model to obtain a deep learning model with the increased quantity of hidden layers; and train (212B) the deep learning model with

the increased quantity of hidden layers based on the current streaming sample data, to obtain a new deep learning model; and

the evaluation unit (306) is configured to determine the latest deep learning model based on a result of performance comparison between the new deep learning

5 model and the current deep learning model.

~~1210~~. The apparatus (300) according to claim ~~119~~, wherein the parameters of the shallow learning model are used as initialization parameters of the deep learning model with the increased quantity of hidden layers.

~~1311~~. The apparatus (300) according to claim ~~119~~ or ~~1210~~, wherein the
10 evaluation unit (306) is specifically configured to:

in response to determining that performance of the new deep learning model is improved compared to the performance of the current deep learning model, use the new deep learning model as the latest deep learning model; or

in response to determining that performance of the new deep learning model is
15 not improved compared to the performance of the current deep learning model, use the current deep learning model as the latest deep learning model.

~~1412~~. The apparatus (300) according to any one of claims ~~119~~ to ~~1311~~, further comprising:

~~A~~a weighting unit (308), configured to weight the latest deep learning model and
20 the shallow learning model to obtain the latest learning model.

~~15. A computing device, comprising:~~

~~at least one processor; and~~

~~a storage that communicates with the at least one processor, wherein an
executable instruction is stored in the memory, and the at least one processor is
25 enabled to implement the method according to any one of claims 1 to 7 when the
executable instruction is executed by the at least one processor.~~