CLAIMS

1. A method, comprising:

obtaining at least one first sample of an internet activity, each first sample including feature data and a label value, the label value of the first sample corresponding to a predicted value of a first model based on the feature data of the first sample;

obtaining at least one first output value of a second model by separately inputting feature data of each first sample of the at least one first sample into the second model, the second model including a probability function, each first output value being a probability function of the feature data of the first sample of the at least one first sample and indicating a probability that the first sample being selected as a training sample for training the first model;

obtaining a first training sample set from the at least one first sample based on the at least one first output value, a first output value corresponding to a training sample in the first training sample set meeting a first threshold output value;

training the first model using the first training sample set; and

automatically identifying risk events in internet activities using the first model as trained.

2. The method according to claim 1, comprising training the second model through acts including:

obtaining at least one second sample of an internet activity, each second sample including feature data and a label value, the label value of the second sample corresponding to a predicted value of the first model with respect to the feature data of the second sample;

obtaining at least one second output value of the second model by separately inputting feature data of each second sample of the at least one second sample into the second model, each second output value being a probability function of the feature data of the second sample of the at least one second sample and indicating a probability that the second sample being selected as a training sample for training the first model

determining a second training sample set of the first model from the at least one second sample based on the at least one second output value, a second output value corresponding to a training sample in the second training sample set meeting a second threshold output value;

training the first model by using the second training sample set to obtain a trained first model, and obtaining a first predicted loss of the trained first model using a test sample having known feature data and a known label value;

calculating a reward value corresponding to the at least one second output value of the second model based on the first predicted loss; and

training the second model by using a policy gradient algorithm and the feature data of the at least one second sample, each second output value of the second model for each feature data of the at least one second sample, and the reward value.

3. The method according to claim 2, wherein the acts further include after the obtaining the first predicted loss of the trained first model, restoring the first model to include model parameters that exist before the training the first model.

4. The method according to claim 2, wherein the reward value is equal to a difference obtained by subtracting the first predicted loss from an initial predicted loss, and the method further comprises:

after the obtaining the at least one second sample, randomly obtaining an initial training sample set from the at least one second sample; and

training the first model by using the initial training sample set, and obtaining the initial predicted loss of a trained first model based on the test sample.

5. The method according to claim 2, wherein the acts are iterated in multiple iterations of trainings, and the reward value is equal to a difference obtained by subtracting a first predicted loss obtained in a current iteration of the multiple iterations of the trainings from a first predicted loss obtained in a previous iteration of the multiple iterations of the trainings immediately before the current iteration.

6. The method according to claim 2, wherein the at least one first sample is same as the at least one second sample.

7. The method according to claim 1, wherein the first model is an anti-fraud model, the feature data is feature data of a transaction, and the label value indicates whether the transaction is a fraudulent transaction.

8. A computing device, comprising a memory and a processor, the memory storing executable code, and when executing the executable code, the processor implements acts including:

obtaining at least one first sample of an internet activity, each first sample including feature data and a label value, the label value of the first sample corresponding to a predicted value of a first model based on the feature data of the first sample;

obtaining at least one first output value of a second model by separately inputting feature data of each first sample of the at least one first sample into the second model, the second model including a probability function, each first output value being a probability function of the feature data of the first sample of the at least one first sample and indicating a probability that the first sample being selected as a training sample for training the first model;

obtaining a first training sample set from the at least one first sample based on the at least one first output, a first output value corresponding to a training sample in the first training sample set meeting a first threshold output value;

training the first model using the first training sample set; and

automatically identifying risk events in internet activities using the first model as trained.

9. The computing device according to claim 8, wherein the acts comprises training the second model through training actions including:

obtaining at least one second sample of an internet activity, each second sample including feature data and a label value, the label value of the second sample corresponding to a predicted value of the first model with respect to the feature data of the second sample;

obtaining at least one second output value of the second model by separately inputting feature data of each second sample of the at least one second sample into the second model, each second output value being a probability function of the feature data of the second sample of the

at least one second sample and indicating a probability that the second sample being selected as a training sample for training the first model;

determining a second training sample set of the first model from the at least one second sample based on the at least one second output value, a second output value corresponding to a training sample in the second training sample set meeting a second threshold output value;

training the first model by using the second training sample set to obtain a trained first model, and obtaining a first predicted loss of the trained first model using a test sample having known feature data and a known label value;

calculating a reward value corresponding to each of the at least one second output value of the second model based on the first predicted loss; and

training the second model by using a policy gradient algorithm and the feature data of the at least one second sample, each second output value of the second model for each feature data of the at least one second sample, and the reward value.

10. The computing device according to claim 9, wherein the training actions further include after the obtaining the first predicted loss of the trained first model, restoring the first model to include model parameters that exist before the training the first model.

11. The computing device according to claim 9, wherein the reward value is equal to a difference obtained by subtracting the first predicted loss from an initial predicted loss, and the training actions further include:

after the obtaining the at least one second sample, randomly obtaining an initial training sample set from the at least one second sample; and

training the first model by using the initial training sample set, and obtaining the initial predicted loss of a trained first model based on the test sample.

12. The computing device according to claim 9, wherein the training actions are iterated in multiple iterations of trainings, and the reward value is equal to a difference obtained by subtracting a first predicted loss obtained in a current iteration of the multiple iterations of

trainings from a first predicted loss obtained in a previous iteration of the multiple iterations of trainings immediately before the current iteration.

13. The computing device according to claim 8, wherein the first model is an anti-fraud model, the feature data is feature data of a transaction, and the label value indicates whether the transaction is a fraudulent transaction.

14. A non-transitory storage medium having executable instructions stored thereon, the executable instructions, when executed by a processor, enable the processor to implement acts comprising:

obtaining at least one first sample of an internet activity, each first sample including feature data and a label value, the label value of the first sample corresponding to a predicted value of a first model based on the feature data of the first sample;

obtaining at least one first output value of a second model by separately inputting feature data of each first sample of the at least one first sample into the second model, the second model including a probability function, each first output value being a probability function of the feature data of the first sample of the at least one first sample and indicating a probability that the first sample being selected as a training sample for training the first model;

obtaining a first training sample set from the at least one first sample based on the at least one first output value, a first output value corresponding to a training sample in the first training sample set meeting a first threshold output value;

training the first model using the first training sample set; and

automatically identifying risk events in internet activities using the first model as trained.

15. The non-transitory storage medium according to claim 14, wherein the acts include training the second model through training actions including:

obtaining at least one second sample of an internet activity, each second sample including feature data and a label value, the label value of the second sample corresponding to a predicted value of the first model with respect to the feature data of the second sample;

obtaining at least one second output value of the second model by separately inputting feature data of each second sample of the at least one second sample into the second model, each second output value being a probability function of the feature data of the second sample of the at least one second sample and indicating a probability that the second sample being selected as a training sample for training the first model

determining a second training sample set of the first model from the at least one second sample based on the at least one second output value, a second output value corresponding to a training sample in the second training sample set meeting a second threshold output value;

training the first model by using the second training sample set to obtain a trained first model, and obtaining a first predicted loss of the trained first model using a test sample having known feature data and a known label value;

calculating a reward value corresponding to the at least one second output value of the second model based on the first predicted loss; and

training the second model by using a policy gradient algorithm and the feature data of the at least one second sample, each second output value of the second model for each feature data of the at least one second sample, and the reward value.

16. The non-transitory storage medium according to claim 15, wherein the training actions include after the obtaining the first predicted loss of the trained first model, restoring the first model to include model parameters that exist before the training the first model.

17. The non-transitory storage medium according to claim 15, wherein the reward value is equal to a difference obtained by subtracting the first predicted loss from an initial predicted loss, and the training actions include:

after the obtaining the at least one second sample, randomly obtaining an initial training sample set from the at least one second sample; and

training the first model by using the initial training sample set, and obtaining the initial predicted loss of a trained first model based on the test sample.

18. The non-transitory storage medium according to claim 15, wherein the training actions are iterated in multiple iterations of trainings, and the reward value is equal to a difference obtained by subtracting a first predicted loss obtained in a current iteration of the multiple iterations of the trainings from a first predicted loss obtained in a previous iteration of the multiple iterations of the trainings immediately before the current iteration.

19. The non-transitory storage medium according to claim 15, wherein the at least one first sample is same as the at least one second sample.

20. The non-transitory storage medium according to claim 14, wherein the first model is an anti-fraud model, the feature data is feature data of a transaction, and the label value indicates whether the transaction is a fraudulent transaction.