CLAIMS

What is claimed is:

1. A recommendation method, wherein the method comprises:

obtaining a target feature vector, wherein the target feature vector is obtained by performing feature extraction on attribute information of both a target user and a target article;

processing the target feature vector by using a recommendation model, to obtain recommendation information, wherein the recommendation information indicates a probability that the target user selects the target article, the recommendation model comprises a cross network (cross network), a deep network (deep network), and a target network, the cross network comprises a first cross layer (cross layer) and a second cross layer, and the deep network comprises a first deep layer (deep layer) and a second deep layer; the target network is used to perform fusion processing on a first intermediate output that is output by the first cross layer and a second intermediate output that is output by the first deep layer, to obtain a first fusion result, and the target network is further used to: process the first fusion result to obtain a first weight corresponding to the first cross layer and a second weight corresponding to the first deep layer, and weight the first fusion result with the first weight and the second weight separately, to obtain a first intermediate input and a second intermediate input; and the first intermediate input is input data of the second cross layer, and the second intermediate input is input data of the second deep layer; and

when the recommendation information meets a preset condition, determining to recommend the target article to the target user.

2. The method according to claim 1, wherein the fusion processing comprises one of point-wise addition (point-wise addition), Hadamard product, concatenation, and attention mechanism based pooling.

3. The method according to claim 1 or 2, wherein the target network comprises a first feature adaptation network, and the first feature adaptation network is a fully-connected network, a squeeze-and-excitation network (squeeze-and-excitation network), an attention network, an SENet, or a gatenet (gatenet); and

the processing the first fusion result comprises:

processing the first fusion result by using the first feature adaptation network.

4. The method according to any one of claims 1 to 3, wherein the first intermediate output comprises M first feature vectors, the second intermediate output comprises M second feature vectors, the first fusion result comprises M third feature vectors, the first weight comprises a first weight value corresponding to each first feature vector, and the second weight comprises a second

weight value corresponding to each second feature vector; and

the weighting the first fusion result with the first weight and the second weight separately comprises:

weighting each first feature vector of the M third feature vectors with a corresponding first weight value; and

weighting each second feature vector of the M third feature vectors with a corresponding second weight value.

5. The method according to any one of claims 1 to 4, wherein

the cross network further comprises a third cross layer, and the deep network further comprises a third deep layer; the target network is further used to perform fusion processing on a third intermediate output that is output by the second cross layer and a fourth intermediate output that is output by the second deep layer, to obtain a second fusion result, and the target network is further used to: process the second fusion result to obtain a third weight corresponding to the second cross layer and a fourth weight corresponding to the second deep layer, and weight the second fusion result with the third weight and the fourth weight separately, to obtain a third intermediate input corresponding to the second cross layer and a fourth intermediate input corresponding to the second deep layer; and the third cross layer is configured to process the third intermediate input, and the third deep layer is configured to process the fourth intermediate input.

6. The method according to claim 5, wherein the target network comprises a second feature adaptation network, and the second feature adaptation network is a fully-connected network, a squeeze-and-excitation network (squeeze-and-excitation network), an attention network, an SENet, or a gatenet (gatenet); and

the processing the second fusion result comprises:

processing the second fusion result by using the second feature adaptation network.

7. The method according to any one of claims 1 to 6, wherein the target feature vector comprises a first network input and a second network input; and

the obtaining a target feature vector comprises:

obtaining an initial feature vector, wherein the initial feature vector is obtained by performing feature extraction on the attribute information of both the target user and the target article; and

processing the initial feature vector by using a third feature adaptation network, to obtain a fifth weight corresponding to the cross network and a sixth weight corresponding to the deep network, and weighting the initial feature vector with the fifth weight and the sixth weight separately, to obtain the first network input corresponding to the cross network and the second network input corresponding to the deep network, wherein the first network input is used as an input of the cross network, and the second network input is used as an input of the deep network.

8. The method according to claim 7, wherein the third feature adaptation network is a fully-connected network, a squeeze-and-excitation network, an attention network, an SENet or a gatenet.

9. The method according to any one of claims 1 to 8, wherein the user attribute comprises at least one of the following: a gender, an age, an occupation, income, a hobby, and an education degree.

10. The method according to any one of claims 1 to 9, wherein the article attribute comprises at least one of the following: an article name, a developer, an installation package size, a category, and a good rating.

11. A model training method, wherein the method comprises:

obtaining a target feature vector, wherein the target feature vector is obtained by performing feature extraction on attribute information of both a target user and a target article;

processing the target feature vector by using a first recommendation model, to obtain recommendation information, wherein the recommendation information indicates a probability that the target user selects the target article, the recommendation model comprises a cross network (cross network), a deep network (deep network), and a target network, the cross network comprises a first cross layer (cross layer) and a second cross layer, and the deep network comprises a first deep layer (deep layer) and a second deep layer; the target network is used to perform fusion processing on a first intermediate output that is output by the first cross layer and a second intermediate output that is output by the first deep layer, to obtain a first fusion result, and the target network is further used to: process the first fusion result to obtain a first weight corresponding to the first cross layer and a second weight corresponding to the first deep layer, and weight the first fusion result with the first weight and the second weight separately, to obtain a first intermediate input and a second intermediate input; and the first intermediate input is input data of the second cross layer, and the second intermediate input is input data of the second deep layer; and

determining a loss based on the recommendation information and an actual selection result of the target user for the target article, and updating the first recommendation model based on the loss, to obtain a second recommendation model.

12. The method according to claim 11, wherein the fusion processing comprises one of point-wise addition (point-wise addition), Hadamard product, concatenation, and attention mechanism based pooling.

13. A computing device, wherein the computing device comprises a memory and a processor, the memory stores code, and the processor is configured to obtain the code and perform the method according to any one of claims 1 to 12.

14. A computer storage medium, wherein the computer storage medium stores one or more

instructions, and when the instruction is executed by one or more computers, the one or more computers are enabled to perform the method according to any one of claims 1 to 12.

15. A computer program product, comprising code, wherein when the code is executed, the method according to any one of claims 1 to 12 is performed.