CLAIMS

What is claimed is:

1. A recommendation method, comprising:

obtaining a plurality of images, wherein each image comprises one candidate interface and one type of candidate content presented by using the candidate interface;

obtaining image feature data of each image;

predicting, by using a prediction model, a degree of preference of a target user for each image based on user feature data of the target user and the image feature data, wherein input for the prediction model is determined based on the user feature data and the image feature data; and

selecting, based on the degree of preference, a candidate interface and/or candidate content from candidate interfaces and candidate content that are comprised in the plurality of images, to perform recommendation.

2. The method according to claim 1, wherein each image comprises a plurality of regions; and

the image feature data of each image comprises a plurality of local eigenvectors, and each local eigenvector represents one region.

3. The method according to claim 2, wherein the predicting, by using a prediction model, a degree of preference of a target user for each image based on user feature data of the target user and the image feature data comprises:

for each image, obtaining N word vectors based on the candidate content in the image, wherein each word vector represents one word in the candidate content, and N is a positive integer;

for each word vector, calculating, by using a model of an attention mechanism, an attention weight of each of the plurality of local eigenvectors based on the word vector and the plurality of local eigenvectors, wherein the attention weight indicates a degree to which the target user pays attention to a region represented by the local eigenvector when the target user reads a word represented by each word vector;

performing fusion on each word vector and the plurality of local eigenvectors based on the attention weight of each of the plurality of local eigenvectors to obtain a first fusion eigenvector, wherein one first fusion eigenvector is correspondingly obtained for each word vector; and

predicting, by using the prediction model, the degree of preference of the target user for each image based on the user eigenvector and N first fusion eigenvectors corresponding to the N word vectors, wherein input for the prediction model is determined based on the user eigenvector and the N first fusion eigenvectors, and the user eigenvector represents the user feature data of the target user.

4. The method according to claim 3, wherein the predicting, by using the prediction model, the degree of preference of the target user for each image based on the user eigenvector and N first fusion eigenvectors corresponding to the N word vectors comprises:

for each image, processing, by using a model of a self-attention mechanism, the N first fusion eigenvectors corresponding to the N word vectors to obtain N semantic-enhanced eigenvectors, wherein each first fusion eigenvector corresponds to one semantic-enhanced eigenvector; and

predicting, by using the prediction model, the degree of preference of the target user for each image based on the user eigenvector and the N semantic-enhanced eigenvectors, wherein input for the prediction model is determined based on the user eigenvector and the N semantic-enhanced eigenvectors.

5. The method according to claim 4, wherein the predicting, by using the prediction model, the degree of preference of the target user for each image based on the user eigenvector and the N semantic-enhanced eigenvectors comprises:

for each image, performing fusion on the N semantic-enhanced eigenvectors by using a model of an additive attention mechanism to obtain a second fusion eigenvector; and

predicting, by using the prediction model, the degree of preference of the target user for each image based on the user eigenvector and the second fusion eigenvector, wherein input for the prediction model is determined based on the user eigenvector and the second fusion eigenvector.

6. The method according to claim 1, wherein the image feature data of each image comprises a global eigenvector, and the global eigenvector represents the image.

7. The method according to any one of claims 1 to 6, wherein the selecting, based on the degree of preference, a candidate interface and/or candidate content from candidate interfaces and candidate content that are comprised in the plurality of images, to perform recommendation comprises:

selecting, based on the degree of preference, a type of candidate content from the candidate content comprised in the plurality of images as target candidate content; and

selecting, based on the degree of preference, a candidate interface from candidate interfaces of images that comprise the target candidate content as a target candidate interface, to recommend the target candidate content by using the target candidate interface.

8. A training method, comprising:

obtaining a plurality of sample images, wherein each sample image comprises one sample candidate interface and one type of sample candidate content presented by using the sample candidate interface;

obtaining image feature data of each sample image;

predicting, by using a prediction model, a degree of preference of a sample user for each

sample image based on user feature data of the sample user and the image feature data, wherein input for the prediction model is determined based on the user feature data and the image feature data; and

adjusting the prediction model based on the degree of preference and historical click-through data of the sample user for the sample candidate content.

9. The method according to claim 8, wherein each sample image comprises a plurality of regions; and

the image feature data of each sample image comprises a plurality of local eigenvectors, and each local eigenvector represents one region.

10. The method according to claim 9, wherein the predicting, by using a prediction model, a degree of preference of a sample user for each sample image based on user feature data of the sample user and the image feature data comprises:

for each sample image, obtaining N word vectors based on the sample candidate content in each sample image, wherein each word vector represents one word in the sample candidate content, and N is a positive integer;

for each word vector, calculating, by using a model of an attention mechanism, an attention weight of each of the plurality of local eigenvectors based on the word vector and the plurality of local eigenvectors, wherein the attention weight indicates a degree to which the sample user pays attention to a region represented by the local eigenvector when the sample user reads a word represented by each word vector;

performing fusion on each word vector and the plurality of local eigenvectors based on the attention weight of each of the plurality of local eigenvectors to obtain a first fusion eigenvector, wherein one first fusion eigenvector is correspondingly obtained for each word vector; and

predicting, by using the prediction model, the degree of preference of the sample user for each sample image based on the user eigenvector and N first fusion eigenvectors corresponding to the N word vectors, wherein input for the prediction model is determined based on the user eigenvector and the N first fusion eigenvectors, and the user eigenvector represents the user feature data of the sample user.

11. A computer device, comprising a memory and a processor, wherein the memory is configured to store computer-readable instructions, and the processor is configured to read the computer-readable instructions and implement the method according to any one of claims 1 to 7.

12. A training device, comprising a memory and a processor, wherein the memory is configured to store computer-readable instructions, and the processor is configured to read the computer-readable instructions and implement the method according to any one of claims 8 to 10.

13. A computer storage medium, wherein the computer storage medium stores computer-

readable instructions, and when the computer-readable instructions are executed by a processor, the method according to any one of claims 1 to 10 is implemented.

14. A computer program product, wherein the computer program product stores computer-readable instructions, and when the computer-readable instructions are executed by a processor, the method according to any one of claims 1 to 10 is implemented.

15. A recommendation system, comprising a terminal device and a server, wherein

the server is configured to perform the method according to any one of claims 1 to 7; and

the terminal device is configured to: receive target candidate content and metadata of a target candidate interface from the server; and

display the target candidate interface based on the metadata, and recommend the target candidate content to a target user by using the target candidate interface.